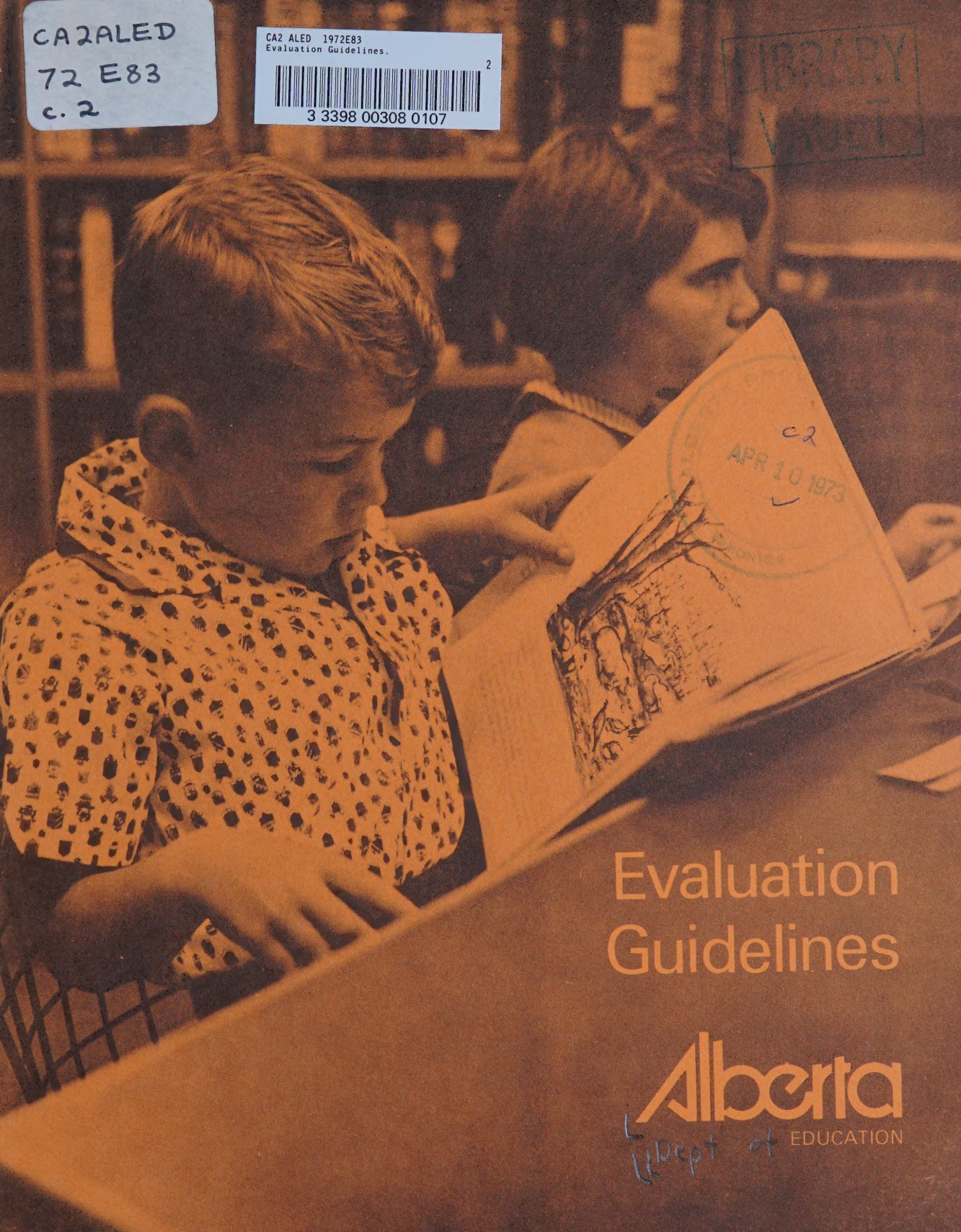
CA2ALED
72 E83
c. 2

CA2 ALED 1972E83
Evaluation Guidelines.
2
3 3398 00308 0107

LIBRARY
VAULT

C2
APR 10 1973



Evaluation Guidelines

Alberta
EDUCATION

Dept of

COVER: Information Canada
Photo by R. Gilliat

Digitized by the Internet Archive
in 2025 with funding from
Legislative Assembly of Alberta - Alberta Legislature Library

Digitized by the Internet Archive
in 2025 with funding from
Legislative Assembly of Alberta - Alberta Legislature Library

https://archive.org/details/ableg_33398003080107

Evaluation Guidelines

by

H. C. Rhodes
Education Consultant
C. A. H. Lomas
Econometrician

edited by
J. Collins Meek
Special Projects Consultant

Alberta
EDUCATION
Executive Building
10105 - 109 Street
Edmonton, Alberta

1972

Preface

This manual is addressed to all personnel within a school jurisdiction who are directly or indirectly involved in the educational process. In particular, it is addressed to those personnel who wish to conduct an evaluation of any activity or group of activities, from the most minute (one class period) to the most massive (the entire school system).

It is thought that this publication will be useful for teachers (in evaluating their own activities), for subject area specialists (in evaluating an on-going program of studies), for program planners (in evaluating a new program or project), for curriculum development personnel (in evaluation of new curricula), and consultants of many kinds. The guidelines and checklists, outlined in this document, while directed at instruction, can also be adapted for use in evaluating such things as an in-service project for teachers, or the administrative organization of a school.

Evaluation may be said to be a series of activities which deliver useful information to decision-makers (be they students, parents, teachers, consultants, or administrators) so that improved educational and supporting programs may result. The information derived from a systematic evaluation, in addition to aiding educational decision-makers at all levels, provides an objective basis for effective consultation on, and discussion of, educational activities between and among the various publics and groups. Hence, evaluation is seen as useful to all individuals involved in new projects or on-going programs, whether as administrators or on a day-to-day or even occasional basis.

In sum, the question may be posed: which people, or groups of people, require information that will permit them to make more accurate assessments or judgments relative to educational or supporting programs? This document attempts to provide guidelines for evaluations which will deliver the required information to the identified persons.

– W.R. Duke
Acting Director, Field Services
Alberta Department of Education

TABLE OF CONTENTS

INTRODUCTION 1

I. EVALUATION: ITS DEFINITION AND PURPOSES 3
- A. Evaluation Defined 3
- B. Evaluation and Research 7
- C. General Guidelines: A Checklist 8

II. GUIDELINES FOR EVALUATION 12
- A. Four Stages of Evaluation: Four Types of Decisions 12
- B. Overview of Four Stages of Evaluation 12
- C. Needs Assessment 15
- D. Program Planning 16
- E. Procedures Evaluation 24
- F. Outcome Evaluation 27
- G. Summary 27

III. OBJECTIVES 32
- A. Purposes and Use of Objectives 32
- B. Types of Objectives 33
- C. How to Write Terminal Objectives 40
- D. Evaluating the Terminal Objectives 44
- E. How to Write Enabling Objectives 44
- F. Evaluation of Enabling Objectives 46
- G. Summary 46

IV. INSTRUMENTATION AND INFORMATION 49
- A. Relationship Between Objectives and Information 49
- B. Classification of Information 49
- C. How to Develop Locally-Produced Instruments 55
- D. Collection and Processing of Information 62
- E. Example of an Information Collection Schedule 68
- F. Summary 72

V. BUDGETING FOR THE PROGRAM AND ITS EVALUATION 76

VI. SOURCES OF ASSISTANCE 78

Extra Copies of Checklists 81

LIST OF FIGURES

1.1 Evaluation Checklist 9
2.1 Program Evaluation Sequence for New Programs and Existing Programs 13
2.2 Recycling Options in the Evaluation Sequence 13
2.3 Flexible Features of the Evaluation Sequence 14
2.4 Establishing Objectives Through Discrepancy Information 15
2.5 Needs Assessment Checklist 17
2.6 Describing School Climate: Information Sources and Observation Procedures 18
2.7 The Relationship Between Procedures Evaluation and Outcome Evaluation 21
2.8 Development and Use of Objectives by Evaluation State 24
2.9 Program Planning Checklist 25
2.10 Decision Areas and Evaluation Activities 26
2.11 Procedures Evaluation Checklist 28
2.12 Outcome Evaluation Checklist 29
3.1 Levels of an Enabling Objective 35
3.2 Levels of a Terminal Objective 37
3.3 Objectives Classification in Two Dimensions 39
3.4 Objectives Classification in Three Dimensions 40
3.5 Evaluative Criteria for Terminal Objectives 45
3.6 Evaluative Criteria for Enabling Objectives 47
4.1 Two Dimensions of the Classification of Information 52
4.2 Source of Instrument Dimension 54
4.3 Three Dimensions of the Classification of Information 52
4.4 Essay Tests vs. Objective Tests 57
4.5 Test Blueprint - Classification of Test Items 58
4.6 Respondent Sampling 64
4.7 Respondent-Student Sampling 66
4.8 Item Sampling 67
5.1 Budgeting for the Resources Required 77
6.1 Common Sources of Assistance 79

Introduction

Evaluation activities are an attempt to determine how well educational objectives are achieved. Thus, objectives are the principle guidelines for evaluation.

The writers suggest there are two kinds of objectives: **terminal objectives** and **enabling objectives**. Terminal objectives provide the main guideline during Outcome Evaluation. During a type of evaluation termed Procedures Evaluation, enabling objectives provide the focus. The development and use of terminal objectives and of enabling objectives constitute the fabric with which the discussions in the various chapters are woven.

There are two kinds of objectives and two kinds of evaluation. The central importance of this perspective warrants a brief discussion in this introduction.

Terminal objectives are phrased as statements of intentions. They are the ends one expects to achieve. Outcome evaluation requires and uses terminal objectives. Outcome evaluation is described as those activities which assist in assessing the degree to which the terminal objectives have been achieved.

The ends (terminal objectives) are achieved through various arrangements of ways and means. Descriptions of intended ways and means are defined as enabling objectives. Evaluation which focuses on information and judgment concerning ways and means is termed, in this publication, Procedures Evaluation. The guideline and the vehicle for Procedures Evaluation are the enabling objectives. Procedures evaluation simply refers to activities undertaken to ascertain how well the enabling objectives have been achieved.

Although evaluators may opt for one or the other of outcome or procedures evaluation, in one sense the separation is artificial. Procedures evaluation is always related, at least by implication, to terminal objectives. This is so since ways and means--the procedures employed--are expected to achieve certain specified ends. Thus, although the terminal objectives and outcome evaluation may be given short shrift in a particular evaluation they are nevertheless an influence. One does not conduct a Procedures Evaluation without reference, however indirectly, to outcome concerns.

In contrast, outcome evaluation may be undertaken without reference to procedures. Enabling objectives in this instance may be ignored. The purpose of the evaluation provides the guideline here. If information is required on which to base decisions about terminating, adopting or expanding a program, outcome evaluation is appropriate. If information is required on which to base decisions about improving a program, then procedures evaluation is the most helpful. If information is required on which to decide the impact of certain factors and procedures on outcomes, and to provide a basis for decisions concerning adoption, expansion, termination *and* improvement, then both procedures evaluation and outcome evaluation are necessary.

A complete evaluation requires both approaches. Procedures evaluation assumes outcome information is available concerning actual or expected achievement of terminal objectives.

To complete this introduction, brief mention is made of the purposes and content of each chapter. Chapter I provides a general overview of evaluation, its definition and purposes. Key terms in the definition--systematic collection of information, decision-making, and objectives--are examined. To provide guidelines for program implementers and program evaluators, Chapter II describes the four stages of evaluation. Chapter III discusses how to write and how to use goals and objectives. To increase the specificity of guidelines as

they relate to objectives the two main classifications, terminal and enabling objectives, are further sub-divided. Determination of how well the various objectives are acheived requires information. Information--its sources, collection and processing--is the topic treated in Chapter IV. Chapter V alerts the reader to the need for budgeting. Considered here are the constraints which are operating and the resources available both for program implementation and program evaluation. Chapter VI refers briefly to sources and use of various forms of assistance.

I

Evaluation: Its Definition and Purposes

A. EVALUATION DEFINED

What is evaluation? Why evaluate? Answers to these questions will be provided by firstly defining the activity and, secondly, by discussing implications of terms used in the definition. The definition, and the subsequent discussion should provide a general orientation to those participating in an evaluation. The definition may thus be considered a major guideline. As a general frame of reference it points direction, suggests appropriate activities, and contributes to consistency.

> *The process of evaluation may be defined as systematically collecting and processing information for the purpose of decision-making. The information gathered, and the decisions which result, are related directly to objectives.*

Key terms from this definition are discussed in the sections which follow. Discussion of the key terms will provide guidelines for use by evaluators.

1. Systematic Collection of Information

Collecting and processing information in a systematic manner represents a contrast to haphazard, incomplete and unplanned activity. Evaluation, as with most endeavours in education, is not new. The "new look" in evaluation results only from some change in emphasis. One emphasis is the attempt to collect and process information more systematically than in the past. This suggests the need for a plan or an evaluation design.

"Systematic" also implies planning a calendar of activities and events. Plans and dates can, of course, be changed. However, target dates provide benchmarks to facilitate the management of the program and its evaluation.

The validity of the conclusions made and decisions taken are in large measure dependent on the quality of information used. Many constraints are immediately apparent when one considers the information available and its collection. A number of questions must be asked of the collected information. Is it relevant? Will it answer questions posed and relate directly to objectives sought? Is the information reasonably reliable? It should describe a stable situation in which results would not change if collection dates were changed or if another source or other observers were used.

Is the information sufficiently broadly-based to permit valid judgements? Too little data decreases the probability that sufficient knowledge has been acquired to warrant a specific action. The breadth and variety of information sought represents yet another emphasis in contemporary evaluations. This emphasis suggests utilization of information which will supplement the intuitive, impressionistic judgements on which we have often relied. The observations and judgements of competent professionals will continue as our single most valuable source of information. The judgements from this source provide standards which are sometimes not available from other sources. However, total dependence on a single source ignores other possibilities: more reliable sources may be available; greater economy

may be enjoyed through employment of other procedures; and indirect measures may tap information which eludes even the most skilled observers. The richness and variety of information relevant to an evaluation often spans a broad spectrum. Illustrations of the possibilities are provided by several classification schemes. Learner outcomes (framed as terminal objectives), may, for example, be categorized as cognitive, affective and psychomoter. Within each of these classes a number of levels have been defined.

Are resources sufficient to meet the needs of the evaluation? The systematic collection, processing, summarizing, interpreting and reporting of information, and the relating of this information to objectives, are the major evaluation activities. Such activities are demanding of time and resources; hence, an evaluation plan should include a budget for these. A motivated, committed staff which is amenable to change and somewhat empirically-minded is another necessary condition. Finally, leadership and coordination are required.

Can valid judgements be made concerning the information collected? In addition to assurances concerning reliability, relevance and validity of the information, criteria against which results may be compared are required. Unhappy is the evaluator who, after conscientiously collecting and processing information, finds it eludes interpretation. Some of the sources of the required standards are these:

(a) Normative information from similar populations external to the school;
(b) Baseline information descriptive of previous activity and outcomes (pretest data is one example);
(c) Control-group information; and
(d) Specification of intended activities and outcomes prior to installing the programs under examination.

2. Decision-making

Use of the term decision-making in the definition contributes to the clarification of the goal of evaluation. Evaluation provides information for use by decision-makers. Since this is the purpose of evaluation activity, brief attention to types of decisions and the values held by decision-makers is warranted. These and other implications of the use of the term are discussed below.

Decision-making is synonymous with planning. When the orientation is to the future, planning may be the preferred term. For example, when processed, summarized and related to objectives, the descriptive information gathered during the initial phase of an evaluation pinpoints needs. To meet those needs, the staff considers alternative strategies and selects from these. This activity is a basic component in planning.

Inclusion of the terms decision-making or planning in the definition enlarges the concept of evaluation. The terms imply the need to take action. The actions taken are initiated by various types of decisions. Possible decisions are: to continue the present program (and terminate the evaluation); to terminate the program; or to modify the program. Without decisions or plans no purpose has been served. Evaluation, according to this conceptualization, has not occurred. Simply judging the worth or merit of an existing state and neglecting follow-up is incongruent with the expectations held for evaluation.

Evaluation is here viewed as various activities along a time line. The first phase is Needs Assessment. Needs, if established, are requisite to the second phase which is Program Planning. If, all factors considered, a decision is taken to implement the plan, the third type of evaluation, Procedures Evaluation, commences. The information gathered during the progress of the newly-designed program constitutes necessary feedback. Such feedback is

the basis for decisions concerning modifications of objectives, redeployment of resources and changes in any number of strategies. Finally, a last stage may be termed Outcome Evaluation. The evaluation activity at this point is to relate outcomes to objectives. Answers are sought to the question, "Have our intentions been realized?"

At any point on the time line, from Needs Assessment through Planning and Procedures Evaluation to Outcome Evaluation, a number of options are open to school personnel since at any stage they may decide to continue, to modify (to recycle to an earlier stage) or to discontinue the program. The major purpose of the evaluation, that is, to provide a basis for decision-making, has thus been served. Evaluators may elect to suspend judgement and not make a decision during some stage. This suggests that insufficient or inadequate information is available. Gathering, processing and judging (applying standards to) additional information may be necessary. Alternatively, as mentioned above, a decision to terminate may be warranted. During Needs Assessment, for example, the school may decide its objectives are being achieved. If no discrepancies exist between descriptions of the present state or condition and what one aspires to achieve--between "what is" and "what ought to be"--the appropriate decision is to continue the present program and discontinue evaluation.

Any one of the referent groups in a school system (including students) may be classed as decision-makers. Knowledge of the questions, concerns and objectives of each group provide guidelines for all evaluation activities. The values held by these groups may reflect on the credibility of certain types of information. The avowed humanist, for example, may argue that the really important outcomes of education cannot (and should not) be quantified. For this type of decision-maker the results from scales which measure self-actualization or an attitude may be unacceptable.

3. The Role of Objectives

The definition enunciated at the beginning of this publication ended with the statement, *The information gathered, and the decisions which result are related directly to objectives.* The most useful single guideline which may be offered is the admonition to school personnel to refer back constantly to objectives.

Such a guideline is difficult to follow unless objectives are made explicit. (They are, of course, implicit in everything we do). Once the evaluation activity becomes a joint activity with a number of personnel involved, the need for explicit statements increases. A number of benefits are realized from clear statements of objectives. Ambiguities are removed. All project personnel utilize a common frame of reference. Evaluation activities are more apt to be logically consistent.

With performance standards included in the statements of objectives some indication of the means of measuring achievement is provided. This is necessary to reduce uncertainty concerning success or failure of the program being examined. Project personnel, after energetic and well-intentioned activity, and after completing a description of their activities (including impressionistic judgements), are often plagued with uncertainty concerning effectiveness. The question, "Did we really achieve our purposes?" may be left unanswered. Steps in reducing this uncertainty include developing clear statements of objectives, including performance criteria, and *using* these throughout the program. The point is made again that the information processed, if it is to be useful, must be compared to some standard. The bases for comparison (standards) may be the intentions regarding activities and outcomes established initially. Alternatively, or concurrently, relative standards, such as

normative information or data from comparison groups, may also be used. Standards established at the outset which do not require comparisons with other groups may be called absolute standards. Criteria such as *80 per cent of the class will demonstrate self-dependence by scoring above 60 on the Study Habits Inventory*, are of the absolute type. Relative standards require comparison with similar groups. An objective which reads, *The class mean on the Study Habits Inventory will exceed the mean of the test norm (or the mean of the control group),* contains a relative standard.

Judgements are crucial in evaluations. Explicitly stated standards facilitate making judgements. The required standards may be included in the statements of objectives.

Statements of objectives then, are a very necessary part of the information base required in an evaluation. The most useful statements include performance criteria which are used as standards. For convenience, and as an aid to discussion, objectives may be placed in several categories. Goals, general objectives, and specific objectives subdivide objectives according to level of specificity. Enabling and terminal objectives subdivide objectives on the basis of type and purpose.

All specific objectives are subsumed within a general philosophical orientation or **goal statement**. The school may subscribe, for example, to the intent implicit in a statement that includes the phrase ". . . to meet both the needs of society and of the individual student." Such philosophical goal statements tend to be timeless, long-term and generally acceptable in a specific culture.

The value-laden philosophical statement forms the basis for statements of **general objectives**. As will be illustrated in the chapter on objectives, these general statements of intent clarify the general purposes of the program. They provide a common focus for evaluation participants. Compared to philosophical goal statements, general objectives narrow the focus and increase specificity.

Specific objectives narrow the area of concern still more. The amount of ambiguity in the guidelines is largely a function of the degree to which objectives are specifically stated. The greater the detail in the statement, and the more precise the statement of performance (standards) expected, the greater the guidance provided to program planners, implementers and evaluators. The standards required for valid judgements on which decisions are based, if included in statements of specific objectives, facilitate judging/decision-making.

The usefulness of an objectives-centered approach to evaluation is enhanced if a further classification of objectives is used. General and specific objectives may be categorized as enabling objectives or terminal objectives. Each of these is most appropriately used in one of the evaluation stages.

Enabling objectives are short term objectives which, as the term implies, appear prerequisite to the achievement of desired outcomes. Development of statements of enabling objectives occurs during the Planning stage (they may be modified in the light of later experience). Information is gathered during the next phase when the program plan is implemented (during Procedures Evaluation) to ascertain the degree to which the enabling objectives have been achieved. If a discrepancy between intents and observations occurs, specific modifications to procedures are indicated. Teaching strategies, or classroom organization, for example, may need to be changed.

Achievement of enabling objectives is seen as a means to a end. Since the ends sought in education often are not observable nor quantifiable we are sometimes forced to infer certain outcomes from activities of students, teachers or administrators. Enabling Objectives provide a vehicle for making such inferences. In this instance, evidence that certain enabling objectives have been achieved makes it more likely that related outcomes have also been

the basis for decisions concerning modifications of objectives, redeployment of resources and changes in any number of strategies. Finally, a last stage may be termed Outcome Evaluation. The evaluation activity at this point is to relate outcomes to objectives. Answers are sought to the question, "Have our intentions been realized?"

At any point on the time line, from Needs Assessment through Planning and Procedures Evaluation to Outcome Evaluation, a number of options are open to school personnel since at any stage they may decide to continue, to modify (to recycle to an earlier stage) or to discontinue the program. The major purpose of the evaluation, that is, to provide a basis for decision-making, has thus been served. Evaluators may elect to suspend judgement and not make a decision during some stage. This suggests that insufficient or inadequate information is available. Gathering, processing and judging (applying standards to) additional information may be necessary. Alternatively, as mentioned above, a decision to terminate may be warranted. During Needs Assessment, for example, the school may decide its objectives are being achieved. If no discrepancies exist between descriptions of the present state or condition and what one aspires to achieve--between "what is" and "what ought to be"--the appropriate decision is to continue the present program and discontinue evaluation.

Any one of the referent groups in a school system (including students) may be classed as decision-makers. Knowledge of the questions, concerns and objectives of each group provide guidelines for all evaluation activities. The values held by these groups may reflect on the credibility of certain types of information. The avowed humanist, for example, may argue that the really important outcomes of education cannot (and should not) be quantified. For this type of decision-maker the results from scales which measure self-actualization or an attitude may be unacceptable.

3. The Role of Objectives

The definition enunciated at the beginning of this publication ended with the statement, *The information gathered, and the decisions which result are related directly to objectives.* The most useful single guideline which may be offered is the admonition to school personnel to refer back constantly to objectives.

Such a guideline is difficult to follow unless objectives are made explicit. (They are, of course, implicit in everything we do). Once the evaluation activity becomes a joint activity with a number of personnel involved, the need for explicit statements increases. A number of benefits are realized from clear statements of objectives. Ambiguities are removed. All project personnel utilize a common frame of reference. Evaluation activities are more apt to be logically consistent.

With performance standards included in the statements of objectives some indication of the means of measuring achievement is provided. This is necessary to reduce uncertainty concerning success or failure of the program being examined. Project personnel, after energetic and well-intentioned activity, and after completing a description of their activities (including impressionistic judgements), are often plagued with uncertainty concerning effectiveness. The question, "Did we really achieve our purposes?" may be left unanswered. Steps in reducing this uncertainty include developing clear statements of objectives, including performance criteria, and *using* these throughout the program. The point is made again that the information processed, if it is to be useful, must be compared to some standard. The bases for comparison (standards) may be the intentions regarding activities and outcomes established initially. Alternatively, or concurrently, relative standards, such as

about assisting in decisions of a local nature and this requires adapting the program to meet local needs. On the other hand the researcher is concerned with a much wider audience of decision-maker. He requires results which are generalizable to wider (or at least more clearly defined) sets of circumstances. He also prefers no change in program after its implementation in order to account for the impact of specific factors.

Guidelines to follow in deciding for or against a research design in the evaluation are to choose a research orientation if:

(a) errors in judgments made and the attendant decisions will be costly;
(b) the emphasis is on outcomes, that is, evaluation is undertaken at the completion of the program where the most important decision pertains to widespread adoption;
(c) the emphasis is *not* on the flexibility required for program improvement during the progress of the venture; and
(d) the purpose of the evaluation is to assess a new program, for wide application, during its developmental phase. Stated differently, the purpose here is not to improve a program proven meritorious elsewhere to meet local needs.

Where feasible, elements of traditional research design should always be incorporated as part of evaluation activities. For example, to control for errors attributable to the normal maturation of students, time of year, and the impact of the general culture on changes in pupil behavior the comparative approach is attractive. Here control classes or schools, which are initially described as comparable on the factors measured, are useful. Program personnel may still make changes during the progress of the program (a feature not typical of research design). The program implementers are left free to evaluate formatively--afforded flexibility in making program adjustments–while retaining the benefit of the external, relative standards provided by control-groups. Control-group information allows comparisons during the final (summative) evaluation.

C. GENERAL GUIDELINES: A CHECKLIST

A number of guidelines were explicated or implied in this chapter. To provide further clarification a checklist is provided below. The checklist is designed for use by program developers and implementers. The checklist will also serve those responsible for coordination and management of the evaluation. It provides a means to review the adequacy of planning. Later, it is a convenient tool with which to audit and monitor the progress of the program and its evaluation. Lastly, it summarizes the adequacy of the information base before decisions are taken.

The user may adopt symbols to represent various responses:

\+ = satisfactory, or has received sufficient attention
? = indeterminate, or of questionable use in this evaluation
– = unsatisfactory, or has not yet received attention
0 = not required in this evaluation

The user may use checklist blanks at different times: in planning the evaluation initially, during the evaluation, and as a final check near completion. In planning the evaluation, dates may be used as targets. During the evaluation, dates may be used to record completion of an activity.

Figure 1.1 Evaluation Checklist

EVALUATION CHECKLIST

Checklist used for (check one):

Planning the evaluation design ☐

Recording progress of the evaluation ☐

Reviewing the evaluation ☐

Evaluation Activity or Evaluation Product	Person(s) responsible	+ Satisfactory	? Indeterminate	- Unsatisfactory	0 Not Required	Dates Initiated	Dates Completed
1. Philosophical goal statement(s) concerning program purposes							
2. Statements of general objectives							
3. Descriptions of present state of condition.							
4. Relating descriptive statements to what "ought to be" -- Statements of specific objectives.							
5. Statements of specific objectives in performance terms							
- terminal objectives							
- related enabling objectives							
6. Quality of information							
- relevance and validity							
- reliability							
- breadth and variety							
- credibility							
- accessibility							
7. Performance standards against which information is compared							
- absolute standards							
- relative standards (information from norm or control groups)							

Figure 1.1 (Continued)

	Person(s) responsible	+ Satisfactory	? Indeterminate	− Unsatisfactory	0 Not Required	I Incomplete	Dates Initiated	Dates Completed
8. Evaluation resources								
- budget prepared								
- funding arranged								
- coordinator appointed								
- responsibilities assigned								
- time and/or facilities arranged								
- objectives development								
- and these activities related to information: selecting procedures								
collecting								
processing								
summarizing								
interpreting								
recommending								
reporting								
9. Questions and concerns of decision-makers (list the appropriate groups) Group Question or concern								
10. Decisions taken regarding the program under examination - to continue and adopt								
- to modify								
- to terminate								
- to expand								

D. SUMMARY

The process of evaluation has been defined as the systematic collection and processing of information to provide a basis for decisions, with all activities related directly to objectives. This definition has been discussed, in some detail, to provide a general orientation and a rationale.

Key terms from the definition served to structure the discussion. These are listed below.

Systematic collection and processing of information. Highlighted here was the utilization of evaluation resources according to a plan (within an evaluation design) to accomplish the purposes sought.

Collecting information: concerns and considerations. The kinds and quality of information were discussed. Considered were its relevancy and validity, reliability, breadth and variety.

Using the information: making judgements. The criterion problem, ("compared to what?") was discussed. Without these comparisons interpretation of data is difficult. Relative and absolute criteria (standards) were described and their uses illustrated.

Decision-making. Decisions, if future-oriented, were seen as synonymous with plans. The purpose of evaluation was viewed as provision of an information base for decisions and plans. Evaluation was conceptualized as a sequence of activities and events occurring over time. Events were the decisions to continue, modify or terminate. These could be taken at various points on a time line–during Needs Assessment, Planning, Procedures Evaluation or Outcome Evaluation.

The role of objectives. Reference to objectives during all phases of evaluation constitutes the major guideline to evaluation. Goals, general and specific objectives, were discussed briefly. It was argued that increasing the specificity contributes to structure and consistency in all evaluation activities. The evaluation procedure should match the program's objectives and where the program is designed to make specific changes in pupils, these changes should be the basis for evaluation.

Emphasis on the "new" look in evaluation. The all-pervasive emphasis is the focus on various types of objectives throughout the evaluation. Another emphasis is the attempt to supplement the folklore, tradition and expert opinion on which most decisions have been based in the past. Intuitive judgements of competent observers remains, of course, an important source of information. However, the need to exploit the many and varied sources of other information was seen as obvious. To enhance the utility of all information some emphasis on the search for standards of comparison was noted.

Evaluation and research. Both evaluation and research activities were described as contributing information for use in decisions. Typical evaluation activities produce results less generalizable to other situations and less precise in accounting for specific cause-effect relationships. However, if the major purpose is to decide on program improvements for small groups in a local setting, and if errors are not costly, an evaluation orientation is justified. When expedient, the power and precision of the traditional research designs should always be incorporated in evaluation procedures.

II
Guidelines for Evaluation

A. FOUR STAGES OF EVALUATION: FOUR TYPES OF DECISIONS

The discussion of evaluation procedures is facilitated if the activities are compartmentalized as stages. Elaboration of the kinds of information required in each stage occurs below. Also receiving attention are the comparisons attempted (criterion standards used), judgements made and the decisions taken. The reader is reminded that, during each stage, objectives are of great significance in giving direction to the evaluation activities. Within the frame of reference provided by objectives, information is gathered, processed and summarized to serve the purposes of decision-making. Particular decisions taken at the various stages are (1) Problem selection (during Needs Assessment), (2) Selecting a particular program (Program Planning), (3) Program improvement (Procedures Evaluation), and (4) Program adoption (Outcome Evaluation).

B. OVERVIEW OF FOUR STAGES OF EVALUATION

Evaluation may be conceptualized as a sequence of activities and events. To assist in program and evaluation management these may be placed on a time-line. Activities may be described as the procedures used in collecting, processing, summarizing and making judgments (applying standards) about information. Events are placed as points in time when certain activities commence and end. Events also properly describe decisions taken.

Each stage of evaluation is described somewhat differently according to the activities required and the purposes served by decisions. The starting point for an evaluation depends on whether the information sought concerns a new program or an existing one. Figure 2.1 serves to illustrate.

Although Figure 2.1 appears to illustrate a one-directional process with a single entry point, it should be noted that project personnel may "leap aboard" at the appropriate stage of evaluation. The appropriate stage depends on the priority assigned to the kind of information sought. For example, it is logical to assume that, since information is available on procedures and outcomes in on-going programs, evaluation activity might first emphasize these concerns. In addition, an evaluation could be feasibly recycled to an earlier stage. The recycling option and cyclical nature of the sequence are illustrated in Figure 2.2.

An evaluation which considers, for example, information on transactions and interactions (Procedures Stage) may prompt a decision either to re-plan (return to Program Planning) or to reconsider the needs (objectives) the program purports to meet. In the latter case project personnel have turned back two stages to Needs Assessment.

As will be indicated in the discussion on Needs Assessment, any outcome or procedures information available from on-going programs contributes to an essential first step in Needs Assessment--a description of "what is." This is a description of the present state or condition of the program.

Figure 2.1 Program Evaluation Sequence for New Programs and Existing Programs

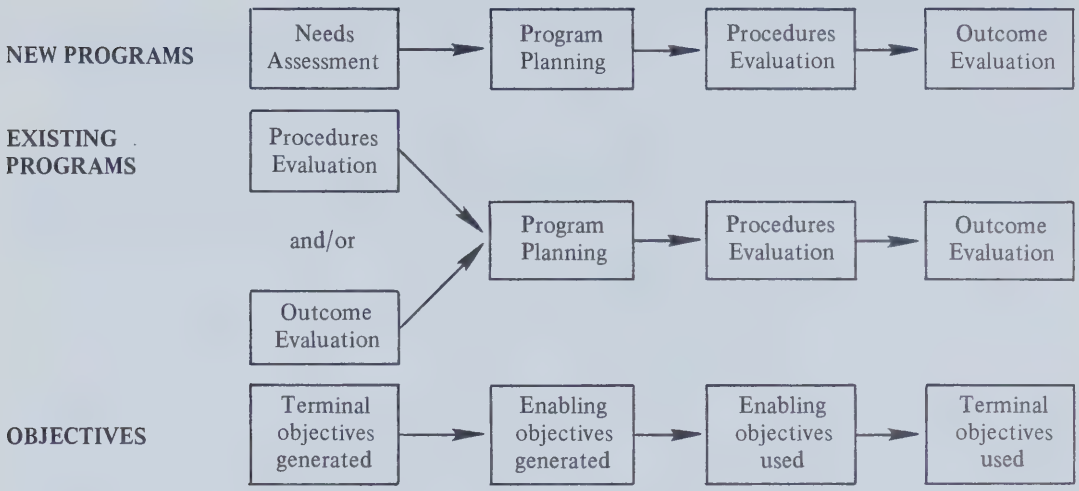
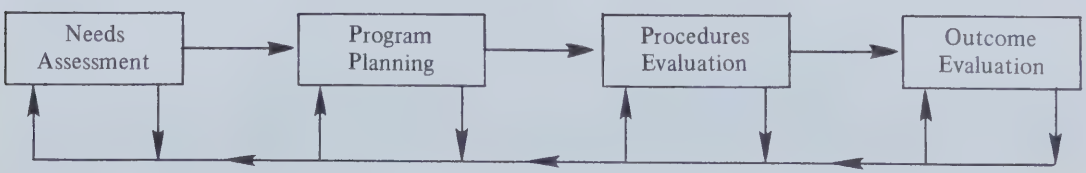
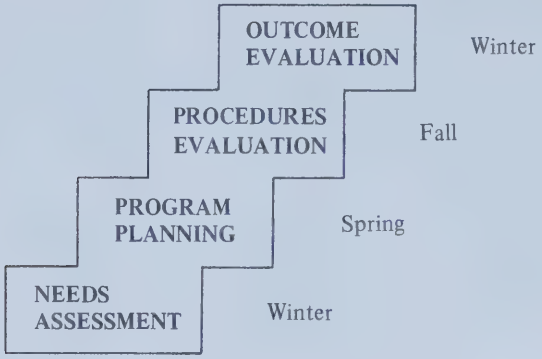


Figure 2.2 Recycling Options in the Evaluation Sequence



A fixed, lock-step procedure should not be implied in the preceding figures. Some of the evaluation activities in the various stages may occur simultaneously. Also, all activities pertinent to one stage need not be completed before the activities in the next begin. This flexibility is portrayed in Figure 2.3.

Figure 2.3 Flexible Features of the Evaluation Sequence



Reference to the four stages of evaluation will be repeated throughout this document. At this time a more detailed treatment is proposed. Use will be made of lists and charts to portray more effectively specific steps which might be followed. The suggestions are few and the lists brief. The reader is expected to supplement to meet the needs of his own evaluation and to effect a transfer to other evaluation contexts.

The sequence developed by the Center for the Study of Evaluation (CSE) at UCLA (1970) has largely been used as the prototype in the model proposed. Features of the Stake (1967, 1968, 1969) model are included as well. The CSE approach has also been embellished with the process-oriented and recycling features of the Provus Discrepancy Evaluation Model (Provus, 1969).

In addition Tyler's work, (Smith:Tyler.1941; Tyler, 1951) emphasizing the dependence of evaluation on objectives, was superimposed on the evaluation model developed for this publication.

The CSE group suggests the existence of five types of decisions required in a normal sequence through which an instructional program passes. The five major decision areas and the kinds of evaluation required for each are listed by the Center for the Study of Evaluation as follows:

DECISION AREA	EVALUATION
1. Problem Selection	Needs Assessment
2. Program Selection	Program Planning
3. Program Operationalization	Implementation Evaluation
4. Program Improvement	Progress Evaluation
5. Program Certification (adoption)	Outcome Evaluation

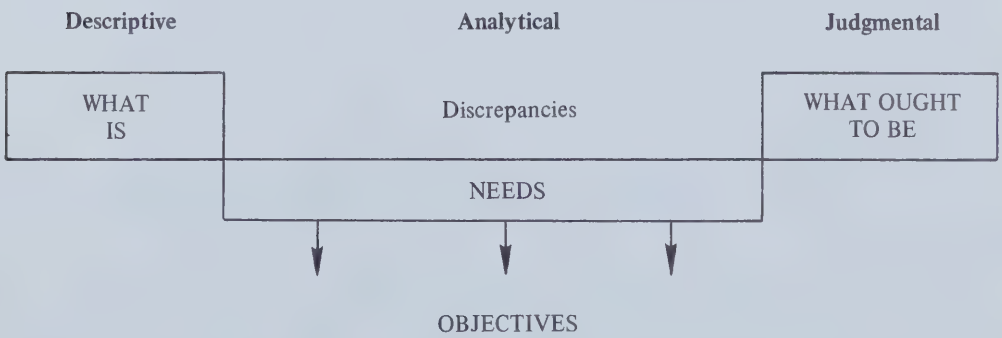
The first two decision areas in the foregoing list are most appropriate if the focus is the school system. The latter three narrow attention to a specific program. In the evaluation model developed in this publication two types of evaluation have been collapsed into one; Implementation Evaluation is simply viewed as an early version of Progress Evaluation. The term applied to the combined phases is Procedures Evaluation.

C. NEEDS ASSESSMENT

Needs assessment requires discrepancy information. It is accomplished by comparing the description of the existing state or condition with what ought to be. To examine this gap two major classes of information are required: (1) the terminal objectives sought, and (2) an adequate description of the current situation in relation to meeting these objectives. A visual representation is noted below.

The category on the left requires adequate information of a descriptive nature. Judgment, of course, is required in selecting the data source and processing techniques. Judgment is crucial in the area pictured at the right where objectives are implied. The discrepancies noted are synonymous with needs. Needs may be re-phrased as objectives. Figure 2.4 illustrates this concept.

Figure 2.4 Establishing Objectives Through Discrepancy Information



Terminal objectives, which are typically expressions of student outcomes, are generated during the Needs Assessment Stage. They are used during the stages which follow, particularly during Outcome Evaluation.

A description of outcomes in the target areas in existing programs is a necessary first step. This descriptive information provides a basis for the next step--judging what ought to be. If differences exist between actual and intended outcomes, these differences may be phrased as terminal objectives. (Enabling objectives, which represent ways and means intended to facilitate achievement of terminal objectives, are developed during the Planning Stage of evaluation. Enabling objectives are always related to specific terminal objectives). Figure 2.5 suggests procedures to be followed during Needs Assessment.

In a particular program now used for illustration--Project School Climate--the *description* of climate was operationally defined by the instrumentation used. A listing of instruments and observation procedures are illustrated in Figure 2.6.

Summaries of results from the administration of the instruments in Figure 2.6 contribute to a description of school climate. The task now facing project personnel is to judge whether the various aspects of climate, as operationally defined by these procedures, is what it ought to be. Again, Figure 2.4 provides a conceptual guide.

For illustration, sample terminal objectives will be developed (with performance standards included) but, to show the relationship to enabling objectives, will be included in the next section. These terminal objectives are generated from the descriptive procedures used in Figure 2.6. Where possible, comparisons are made according to any absolute or relative criteria available. Finally, judgments are made that a discrepancy exists and these are expressed as needs and eventually as terminal objectives. The latter may be communicated in performance terms (e.g. . . . it is expected that, after the one year's experience with the program, student scores on the.....sub-scale of the.....scale will average.....points).

Statements of terminal objectives provide guidance in the planning stage of the evaluation sequence, where, as mentioned previously, the culmination is the development of enabling objectives. Each enabling objective is developed in conjunction with the relevant terminal objectives.

D. PROGRAM PLANNING

Program Planning requires information regarding the probable success of alternative programs in meeting needs. Realistically, this information base must include consideration of constraints. Other factors in the context in which the school operates also need to be considered. Idealism must be tempered with realism. This state of evaluation may eventuate in a decision to implement a particular program–one which has been selected from a number of alternatives.

There are several steps which could be followed during this stage. A short discussion of each is given below.

Constant reference to objectives should occur. Adherence to the common terms of reference, as outlined in statements of philosophical intent, goals, general and specific objectives, facilitates clear communication among all program participants and evaluators. Ambiguities are reduced.

Record the considered views of the staff, central office personnel and outside consultants. A cautionary note is suggested here. Prolonged debate on the outcomes and corresponding ways and means should be avoided.Decisions on designating priorities and selecting enabling and terminal objectives should not be unduly delayed. Features of the Delphi Technique, which has individuals considering issues independently of their peers, has merit (Weaver, 1971). After successive revisions are contributed in writing by individuals, a

Figure 2.5 Needs Assessment Checklist

NEEDS ASSESSMENT CHECKLIST

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
1. Statement of goals						
2. Statement of general objectives						
3. Description of present achievement in the cognitive area Instruments/Indicators/ Observation schedules used						
4. Description of present attitudes (affective areas) Instruments/Indicators/ Observation schedules used						
5. Description of present psychomotor abilities Instruments/Indicators/ Observation schedules used						

continued . . .

Figure 2.5 (Continued)

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
6. Description of discrepancies between "what is" (cognitive, affective and psychomotor levels #3, #4 and #5 above) and "what ought to be" (general objectives)						
DECISION: Which particular discrepancies will become the focus of the program planning stage?						
7. Statements of terminal objectives						
8. Statements of terminal objectives in performance terms						

Figure 2.6 Describing School Climate: Information Sources and Observation Procedures

Instrument or Observation Schedule	Population Sampled
1. Student Beliefs Questionnaire Sub-scales : Friendliness : Honesty : Loyalty : Moral Courage : Responsibility	Grades 5 – 12
2. Children's Philosophy of Human Nature Scale (C-PHN) Sub-scales : Trustworthiness : Strength of Will : Complexity : Altruism : Pessimism	Grades 5 – 12
3. School Morale (SM) Scale Sub-scales. Pupil Morale about : school plant : instruction and instructional materials : administration, regulations and staff : community support and parental involvement : relationships with other students : teacher-student relationships : general feeling about attending school	Grades 5 – 12
4. Grading & Reporting (GR) Survey Pupil beliefs about the purposes of the school as manifested by their opinions about marking and grading practices : values dimension : standards dimension : alternative dimension : system dimension	Grades 7 – 12
5. Junior Index of Motivation (JIM) Scale Pupils' motivation toward school	Grades 7 – 12
6. Self-Concept as a Learner (SCAL) Scale Elementary Secondary	 Grades 3 – 6 Grades 7 – 12

continued . . .

Figure 2.6 (Continued)

7. Self Appraisal Inventory
   Primary — Grades 1–3
   Intermediate — Grades 4–6

   Dimensions are self in relation to
   : family
   : peers
   : scholastic tasks
   : general

8. When do I smile; — Grades 1 – 3
   (General Self Concept)

9. Curriculum Evaluation Student Questionnaire (CESQ) — Grades 10 – 12
   Student opinion about high school programs and programming

10. Student Inventory of Guidance Awareness — Grades 9 – 12
    : Students describe and rate aspects of guidance program including teachers' role.

11. Minnesota Teacher Attitudes Inventory (MTAI) — Teachers
    : Teachers' attitudes toward pupils

12. Flanders Interaction Analysis — Teachers and pupils.
    : Classroom verbal transactions of an affective nature. — Class sessions.

consensus is determined. A forced-choice technique may also provide an expeditious means to accomplish a final consensus.

Review of the research in the area and reports of similar projects elsewhere. Time spent in summarizing the knowledge available in the area may provide invaluable assistance. Some general indices may be consulted which will guide the reader to key articles and references on the topic. Indices which may be consulted are:

(a) Education Index
(b) Encyclopedia of Educational Research
(c) Review of Educational Research
(d) Readers' Guide to Periodical Literature
(e) Library Card Catalogue - Subject Headings
(f) Appropriate Yearbooks of National Associations
(g) ERIC: Research in Education
(h) Current Index to Journals in Education
(i) Dissertation Abstracts

Critical reviews of scales and tests may be found, under the appropriate heading, in O.K. Buros, *Mental Measurements Yearbook* (Nos. 3-7). *Tests in Print* is a companion publication. A review of the literature allows project personnel who have planning responsibilities to profit from the experience of others. The next step also provides this benefit.

Visits to similar projects. Observation of on-going programs contributes to the information the staff requires in selecting between alternatives. It has the additional advantage of assisting the observer through the change sequence. Seeing a program in action apparently has greater credibility than reading about it. Such an experience contributes to a commitment to effect change.

List and consider benefits and disadvantages of alternate programs. Systematic recording of the pros and cons of particular strategies are of obvious benefit.

Select a program. The preceding activities have provided an information base for this decision. At this time a clear delineation of the various aspects of the program is in order.

Define the program. Within the context of the terminal objectives sought, a description of intended use of facilities, resource and personnel provides a standard with which to compare observations made during the next stage--during Procedures Evaluation. A description of intended transactions and interactions (statements of enabling objectives) clarifies the "dynamics" of the plan.

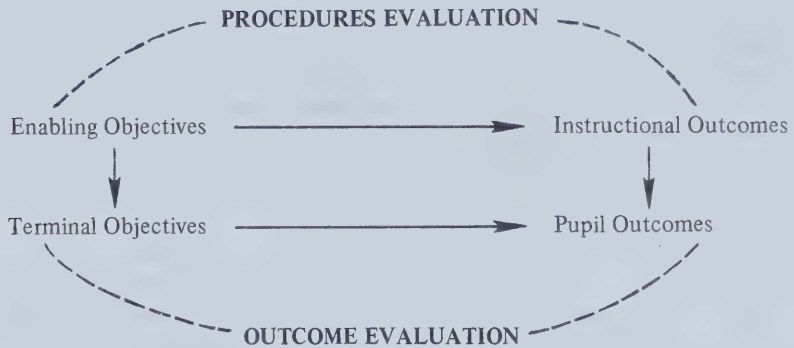
The program definition may be considered in relation to:

(a) Administrators' responsibilities
(b) In-service plans
(c) Organizational arrangements--time schedules, pupil groupings, space assignments, etc.
(d) Materials and media
(e) Teaching strategies and interpersonal relations (transactions)
(f) Diagnosis and feedback
(g) Evaluating and reporting pupil progress

Rephrase the program definition in the form of enabling objectives. Enabling objectives were previously defined as short term objectives considered to be prequisite to achievement of terminal objectives. These statements of intended transactions and plans for use of resources are of use to program evaluators during Procedures Evaluation. They are used as blueprints by program implementers. Enabling objectives focus on instructional procedures

and the teacher. The related terminal objectives focus on pupil outcomes. Figure 2.7 is an attempt to illustrate this essential relationship.

Figure 2.7 The Relationship Between Procedures Evaluation and Outcome Evaluation



Examples from Project School Climate will now be used to illustrate the development and use of terminal objectives. Fragments of the plan selected to achieve the terminal objective may be implied from the statement of the enabling objective. Data from a number of observation procedures used in describing school climate will be used in the first step--describing the present situation. This will be used in the next step--developing performance objectives. Both of these occur during Needs Assessment. Relevant enabling objectives are developed during Program Planning.

EXAMPLE: Using information from *The School Morale Scale*

Description. A low percentage of students (below 50% for most groups) responded favorably to the sub-scale items concerning school plant.

Terminal objective. Next year, at least 65% of all students will respond favorably to items on the school plant sub-scale.

Enabling objectives. (a) Classes will be given recognition for leaving classrooms and laboratories neat and tidy. (b) The caretaking staff will receive social reinforcement for satisfactory work.

EXAMPLE: Using information from *The School Morale Scale*

Description. A high percentage of students agree with item No. 23 of the Instruction and Instructional Materials subscale: *There is too much emphasis on the 'three R's' at this*

school and not enough opportunity for students to develop their own interests. 64.8% of the "all students" group agreed with this statement.

Terminal objective. By January, 1973 less than 40% of the "all student" group will agree with this statement, with no grade responding above 50%.

Enabling objectives. (a) Each course shall be organized so that, at least monthly, students work on tasks related to course objectives but which are of their own choosing. (b) Clubs and project task forces--both long and short-term--will be organized to involve students according to interests.

EXAMPLE: Using results from the Grading and Reporting (GR) Survey

Description. Lack of congruency is noted between teacher and pupil responses on the Values sub-scale. One item from the Values sub-scale reads "The policy of no formal grades is the best policy." Students disagreed with this type of statement more than did teachers.

Terminal objective. After one year's experience with the new pupil progress evaluation program student opinion on the statement will shift significantly toward agreement. (To at least 3.4 on the 6-point rating scale).

Enabling objectives. (1) Student self-evaluation will contribute 50 per cent of the mark assigned by the school in each unit of work. Formal grades at the end of the term will be replaced by pass, no pass and incomplete. (b) Students will contribute to the evaluation criteria to be used assessing progress. (c) Students will consult with teachers concerning tallies in the self-evaluation checklist.

EXAMPLE: Information from item No. 108 of the Curriculum Evaluation Student Questionnaire (CESQ) indicates slight discrepancies with the provincial norm. Grade 11 students in the project school indicate that teachers use a wide variety of methods less frequently. Per cent responses of the local students are compared to the provincial average in the table below. It is agreed that both the local and provincial responses indicate too great an emphasis on lecture.

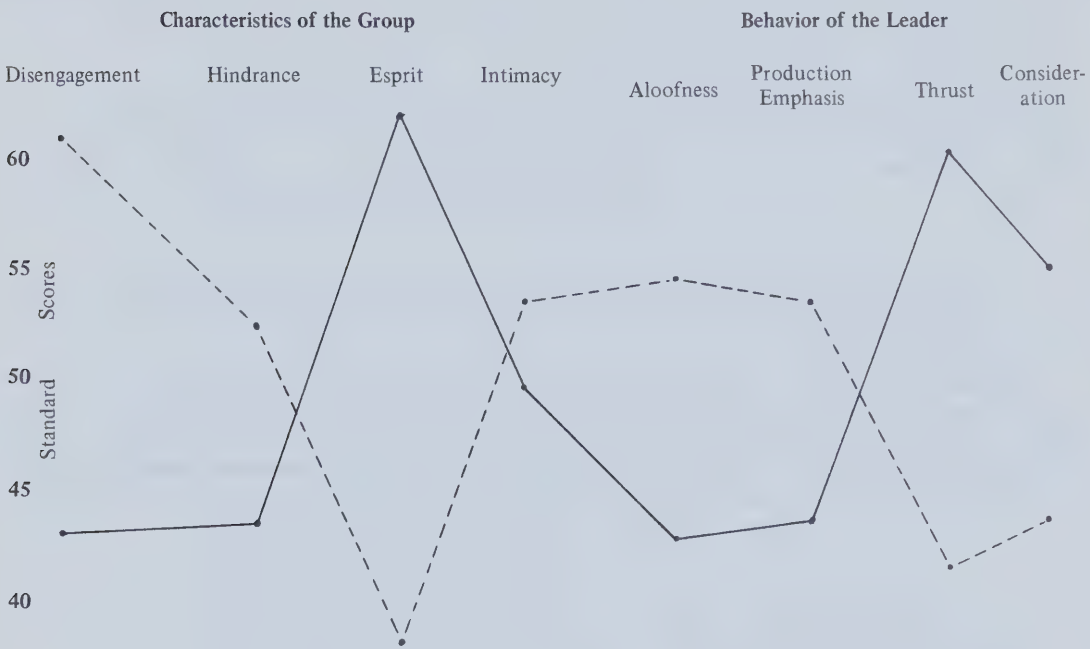
Activity	Local % response	Province %
Teacher lecture	48.5	35.6
Teacher Demonstration	3.0	2.9
Class discussion	33.3	28.5
Small group discussion	6.1	4.6
Individual projects	6.1	4.6
Library research	0.0	18.9
Laboratory experiments	0.0	1.0
Practical activities	0.0	3.3

Terminal objective. To increase pupil perceptions of the frequency with which these instructional methods are used: teacher demonstration, small group discussion, individual projects and library research.

Enabling objective. Teachers record their efforts to increase the frequency of the target strategies.

EXAMPLE: Result from the administration of the Organizational Climate Description Questionnaire (OCDQ) indicate the characteristics of the teacher group and their leader (the principal) approaches the profile of a "closed climate" school.

Description. The *Open* Climate is represented by the solid line. The *Closed* Climate is represented by the dotted line.



Terminal objective. To modify both teacher and principal characteristics, as measured by the OCDQ, to effect a closer approximation to the open climate profile.

Enabling objective. (a) Constitute a principal's cabinet with rotating membership. (b) Cabinet meets weekly with principal and attempts two-way communication between administration and staff. (c) Principal delegates more "paper-work" to school secretary and staff members.

The evaluation stage in which each of terminal and enabling objectives were developed is indicated in the Figure 2.8 below. Later evaluation stages in which each are used are also indicated.

Figure 2.8 Development and Use of Objectives by Evaluation Stage

	Terminal Objectives Developed	Enabling Objectives Developed	Enabling Objectives Used for Evaluation	Terminal Objectives Used for Evaluation
Needs Assessment	X			
Program Planning		X		
Procedures Evaluation			X	
Outcome Evaluation				X

A checklist is provided for use in planning the evaluation design, monitoring the progress of the evaluation and for reviewing the adequacy of the evaluation. This checklist is illustrated in Figure 2.9.

The development of terminal objectives, preferably in performance terms, during the Needs Assessment has provided a basis for planning. The Program Planning stage resulted in the development of enabling objectives designed to facilitate achievement of the terminal objectives. Development of enabling objectives makes possible Procedures Evaluation during the next stage of the sequence--during the implementation and progress of the plan selected.

E. PROCEDURES EVALUATION

If the decision taken as a result of planning is to implement a specific program the next responsibility of evaluation designers is to structure activities for Procedures Evaluation. Examination of the procedures will, as resources permit, provide feedback on how well the enabling objectives have been achieved. A basis is thus provided for decisions to modify the intentions formulated in the plan (to recycle to planning and develop new enabling objectives) and to change the ways and means accordingly. If program improvement has low priority and Procedures Evaluation is ignored, the evaluator has simply gone directly to Outcome Evaluation. If resources are limited, and suitable outcome indicators exist, this may be the most economical strategy. The purpose of the evaluation provides the guideline here. If the purpose is program improvement, Procedures Evaluation is required. If not, and the decisions relate only to adopting or terminating the program or revising outcome expectations, then Outcome Evaluation is necessary and sufficient.

Figure 2.10 below illustrates the type of decision which requires either procedures or outcome evaluation.

Figure 2.9 Program Planning Checklist

PROGRAM PLANNING CHECKLIST

	PERSON(S) RESPONSIBLE	SATISFACTORY.	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
1. Description of present instructional process						
Instruments/Indicators/Observation schedules used						
2. Development of alternative processes or programs						
Review of the literature						
Visits to similar projects						
Judgments of professionals						
Other						
3. Description of factors affecting the choice of a plan						
Budget considerations						
Release time for personnel						
Readiness of professional staff						
Inservice requirements						
Materials available						
Other						

Continued . . .

Figure 2.9 (Continued)

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
4. Description of viable alternatives						
DECISION: Choice of preferred alternative proposal (most viable alternative in terms of benefits and costs)						
5. Description of proposed plan and budget						
6. Statement of enabling objectives*						
Administrative						
Support - personnel						
- facilities						
- materials						
Instructional						
7. Relate enabling objectives to relevant terminal objectives.						

*Instruction includes most teaching activities; support includes activities performed by those who assist the teacher (teacher aide, guidance counselor, librarian, etc.); administrative activities include such things as scheduling, grouping, budgeting, evaluating, etc.

Figure 2.10 Decision Areas and Evaluation Activities

Decision Area	Objectives Used	Evaluation
Program adoption, expansion or termination	Terminal Objectives	Outcome Evaluation
Program improvement and modification	Enabling objectives	Procedures Evaluation

If one purpose of a particular evaluation is to provide information on which to base decisions on how to improve the program then evaluation of procedures is the appropriate activity. The vehicle used is the list of enabling objectives developed during the planning stage. Several types of activity facilitate procedures evaluation:

1. Examination of the degree to which enabling objectives were achieved.
2. Examination of enabling objectives prompted by discrepancy data indicating certain terminal objectives were not achieved.

If certain outcome expectations are not realized, one may search for reasons. This requires documentation concerning the ways and means employed (enabling objectives). Adjustments may then be made involving modifications to enabling objectives–adjustments which are thought to be more productive in achieving the related outcome.

Of course, a weak link exists in educators' knowledge of the mediating factors between means and ends. When ways and means (enabling objectives) are developed and selected we have, in effect, hypothesized that these are our "best bets" for achieving the desired outcomes. When we are disappointed with outcomes a follow-up activity may take the form of revision of the enabling objectives. Here we have developed a new hypothesis concerning the best way to achieve our goals. Such suggestions for new approaches often take the form of recommendations.

We need to be aware of other weaknesses apparent in Procedures Evaluation. We need firstly, to monitor and then to describe procedures. This may be demanding of time and resources. However, monitoring and describing are more successfully achieved than assessing them qualitatively. Documenting the quality of procedures requires a comparison with standards. Very often the only available standards are the judgements of the skilled observer with the requisite competencies which relate to the program.

What steps may be followed in conducting a Procedures Evaluation? The reader is referred to Chapter III where the concept of enabling objectives is treated in more detail. 'Monitoring, observing and describing' suggest information and information-collecting procedures, which are elaborated and illustrated in Chapter IV on Instrumentation and Information. The checklist which follows lists suggested steps intended as guidelines, not prescriptions. Designating persons responsible for each evaluation activity and recording target dates may augment the usefulness of the sequence to those responsible for the management of the program and its evaluation.

In Figure 2.11, step (5), a description of "what is" (recordings of observations) is compared to "what ought to be" (the expectations made explicit in the statement of the enabling objective, as in Needs Assessment. Differences between actual and intended activities constitute discrepancies. Discrepancy information provides a basis for revising activities to bring them into line with objectives, and/or revising the enabling objectives to further improve the program. Such revisions are recommended in order that improvement may occur.

Initiation of Procedures Evaluation may be prompted by awareness of discrepancies between intended and actual outcomes. Outcome evaluation receives attention in the next section.

F. OUTCOME EVALUATION

Much of the information required in Outcome Evaluation will have already been developed during needs assessment. The statements of terminal objectives were generated and the information sources and procedures were selected during the first stage of the evaluation. During Needs Assessment specific information-collecting procedures were employed to, firstly, describe the situation at that time and secondly, to provide the basis for developing statements of terminal objectives. These information gathering procedures may be re-employed in outcome evaluation.

Outcome evaluation requires discrepancy information. A comparison is made between expected achievement (terminal objectives) and recorded observations of actual achievement. Lack of congruency (a discrepancy) between these may provide a basis for modifying, terminating or adopting the program.

Once again, in matters pertaining to objectives the reader is referred to Chapter III. Sources and collection of information receive attention in Chapter IV. A sequence which may provide guidelines to those responsible for program management or for outcome evaluation is outlined in Figure 2.12.

G. SUMMARY

In this chapter evaluation was conceptualized as occurring in four stages. When viewed as a sequence, evaluation activity (and the program being examined) could be terminated during any stage. Alternatively the stages were viewed as somewhat independent and one stage could be selected for attention. However, this was said to represent an incomplete evaluation which avoided follow-up. Guidelines were discussed and listed for each of Needs Assessment, Program Planning, Procedures Evaluation and Outcome Evaluation. In Needs Assessment, information descriptive of existing outcomes was compared to intended outcomes. Discrepancies resulting from this analysis were considered synonymous with needs and needs in turn formed the basis for formulating statements of terminal objectives. Ways and means of achieving terminal objectives were labelled enabling objectives and these were developed during the Program Planning stage. During the third stage--Procedures Evaluation--information was gathered to ascertain how well the enabling objectives were achieved. Finally, guidelines were established for Outcome Evaluation. This fourth stage utilized discrepancy information--intentions compared to actual outcomes--to determine the success in achieving terminal objectives.

Figure 2.11 Procedures Evaluation Checklist

PROCEDURES EVALUATION CHECKLIST

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
1. State the enabling objectives						
2. Relate (1) to relevant terminal objectives						
3. Observe (monitor) the implemented instructional procedures Instruments/Indicators/ Observation schedules used						
4. Describe (record) observations of the procedures						
5. Description of discrepancies between observed procedures and enabling objectives						
6. Apply standards (make judgments) concerning discrepancies						
7. Summarize information from 1 - 6						
DECISION: Should the program be continued? Terminated? Modified? 8. If the decision is to modify-- suggest revisions to enabling objectives (formulate recommendations)						
9. Prepare the report for the procedures evaluation.						

Figure 2.12 Outcome Evaluation Checklist

OUTCOME EVALUATION CHECKLIST

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
For each group of terminal objectives:						
1. Review the statements of terminal objectives						
2. Observe (monitor) the relevant outcome behaviors:						
of achievement in the cognitive area						
Instruments/Indicators/Observation schedules used						
of achievement in the affective area						
Instruments/Indicators/Observation schedules used						
of achievement in the psychomotor area						
Instruments/Indicators/Observation schedules used						

Continued . . .

Figure 2.12 (Continued)

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
3. Describe (record) observations of outcome behaviors						
4. Describe discrepancies between observed outcomes and the terminal objectives						
5. Apply standards (make judgments) concerning discrepancies						
6. Summarize information from 1 - 5						
DECISION: To terminate? To continue? To modify? 7. If decision is to modify suggest revisions to terminal objectives (formulate recommendations), or if information is available on the related enabling objectives suggest revisions (formulate recommendations). If information is unavailable re: enabling objectives initiate procedures evaluation						
8. Prepare the report for the outcome evaluation						

REFERENCES CHAPTER II

Alberta Department of Education, *Responding to Change: A Handbook for Teachers of Secondary Social Studies* Edmonton Department of Education, 1972, Chapter 7-8.

Armstrong, Robert J. and others, *The Development and Evaluation of Behavioral Objectives,* Worthington, Ohio; Chas. A. Jones Publishing Co., 1970.

Center for the Study of Evaluation, "A framework to guide evaluation product development," *UCLA Evaluation Comment,* 2(3), September 1970, 1-4.

Provus, Malcolm, "Evaluation of on-going programs in the public school system," in Ralph Tyler (editor) *Educational evaluation: New roles, new means,* The 68th yearbook of the National Society for the Study of Education, Part III. Chicago, University of Chicago Press, 1969.

Smith, E. R. and Tyler, Ralph W., *Appraising and Recording Student Progress*, New York: Harper and Row 1942.

Stake, Robert E., "The countenance of educational evaluation," *Teachers College Record,* 68(7), April, 1967, 523-540.

Stake, Robert E., "Testing in the evaluation of curriculum development," *Review of Educational Research,* 38, 1968, 77-84.

Tyler, Ralph W., "The functions of measurement in improving instruction," in Lindquist, E. G. (editor), *Educational Measurement*, Washington, D. C.: American Council on Education, 1951.

Weaver, Timothy W., "The Delphi forecasting technique," *Phi Delta Kappan,* 52(5), January, 1971, 267-271.

Armstrong, Harold R., *A Teacher's Guide to Teaching-Performance Evaluation.* Worthington, Ohio: School Management Institute, 1972. (Address: School Management Institue, 6800 High Street, Worthington, Ohio, 43085).

III
Objectives

A. PURPOSE AND USE OF OBJECTIVES

If you don't know where you're going, you're likely to end up somewhere else.

The first step in the evaluative process consists of deciding what we want to evaluate. What aspects of the educational system do we want to improve? Where are the greatest deficiencies? Webster defines deficiency as "not up to a normal standard or complement." In other words, in order to discover deficiencies and areas which need improvement, a normal standard is required. Objectives may define such normal standards. Inclusion of performance criteria in the statements of objectives accomplishes this definition.

Every activity carried out in the educational system is undertaken to achieve some purpose. However, the purpose of various activities is seldom explicitly stated. It can be argued that good teaching can occur without explicitly defined terminal or enabling objectives. (However, objectives can be implied from observations of activities). The quality of an evaluation is directly related to the explicitness with which the purposes of the educational activities are expressed. In evaluation, objectives serve as guidelines for collecting information. There is no sense in collecting information relating to a student's knowledge of math if your objectives relate to his peer-group relations. Thus, objectives guide the information collection process. In addition, an evaluator needs a standard against which to judge the outcomes of an educational program. Objectives serve as one form of standard. This, of course, implies that objectives must be well defined. If this definition has not taken place, then the evaluator (who may not be the objective-writer) may be unsure as to the meaning of the objective. The results of the evaluation may differ from one evaluator to another, leading to unreliable evaluation.

In addition, the act of defining terminal objectives can be useful, even if these objectives are not used for evaluative purposes. Definition of terminal objectives entails a searching process which answers several questions. What are we really asking students to achieve? What particular knowledge, skills and attitudes do we want students to learn? What is the purpose of education? of teaching science? of Physics 30? Answers to questions such as these form the basis for delineation of terminal objectives. This questioning process forces a re-thinking of the purposes behind educational activities. It promotes change, and uncovers discrepancies. If the focus is the student (our ultimate concern), structure and organization of the system is down-played and set in perspective.

In addition, terminal objectives define and provide the prerequisite to learning more accurately, thus aiding curriculum design. The use of well-stated terminal and enabling objectives eliminates many problems in that they help the teacher plan what to teach and when to teach it. Use of terminal objectives alone does not lead to highly structured instruction since objectives do not specify the order or pace at which they will be covered. Enabling objectives, however, do provide structure for the instructional process.

As well as aiding program evaluation, terminal objectives facilitate student evaluation. With the use of objectives, selection and development of valid achievement measures is facilitated. Terminal objectives help teachers determine content validity--the extent to

which a test or observation device measures a representative sample of the subject matter content and the behavioral changes under consideration.

To summarize, objectives for educational activities have many uses. They facilitate the planning process by defining what the system, or a particular program, is trying to achieve. Specifically stated objectives, which include performance criteria, facilitate the evaluative process by providing a standard against which the activities of the system can be judged. Objectives contribute guidelines which direct the educational effort. They are useful both in developing a new program, and in evaluating both new and on-going programs.

Although goals and objectives play a major role in the evaluative process, it must be remembered that the objectives themselves must be evaluated. A goal is timeless, in that it represents a broad purpose or aim. An objective, however, is a highly specific statement of intent. It is probable that a system's goals will not change greatly over time. Specific objectives, however, can change frequently and, therefore, must be evaluated at regular intervals.

A terminal objective which is oriented to a specific grade may not need much revision from year to year. For example, an objective for a particular science class may be: *At the end of the unit, the students will demonstrate, on a teacher-made test, that they know the parts of the human eye, and the relationship between them, with 65% mastery.* When each group of students achieves this objective they move on to the objectives of the next unit. However, the objective remains, ready for the next group of students.

On the other hand, a program may be designed for a particular group of students, in which case, the objectives may change over time. For example, objectives which have been defined to relate to a group of hearing-handicapped students would not be appropriate for the next class, which may consist of physically-disabled children. Or, a remedial reading program would probably have sets of terminal objectives for each student. As the student achieved the objectives set for him, revised objectives would be needed. Since each student would have unique problems, defining one set of objectives for the entire group would be a difficult task. It can be seen that student-specific terminal objectives would need evaluation much more frequently than would subject or grade-specific terminal objectives.

B. TYPES OF OBJECTIVES

Objectives can be classified in many ways. Several of these classifications are given below. Each type of classification illuminates a certain dimension of the objective. They can be used in conjunction to give a full and accurate description of the objective. This classification scheme aids the evaluation of objectives. The three major classification dimensions are: (1) enabling vs. terminal objectives; (2) specific vs. general objectives; and (3) short-range vs. long-range objectives.

1. Enabling vs. Terminal Objectives

There are several major differences between enabling and terminal objectives:

Enabling	Terminal
– relate to ways and means employed in: instruction, support, and administration	– relate to student achievement of: cognitive, affective, and psychomotor skills

Enabling	Terminal
– developed during Program Planning Phase	– developed during Needs Assessment Phase
– used during Procedures Evaluation	– used during Outcome Evaluation

Enabling objectives provide a means for evaluating the process of instruction. It is assumed that certain instructional settings lead to particular student outcomes. Or, to reverse the statement, desirable student outcomes will not be achieved if certain learning situations are not available. Given that this is true, then there should be available a means for evaluating such learning situations and instructional settings. Enabling objectives provide a means for such evaluation.

In one sense, evaluation of the instructional process is no different than evaluation of student outcomes, in that both types of evaluation must, by definition, be based on objectives. However, the differences appear in the content of the objectives (instructional procedures vs. student outcomes) and the phase of evaluation during which these objectives are set (program planning vs. needs assessment.)

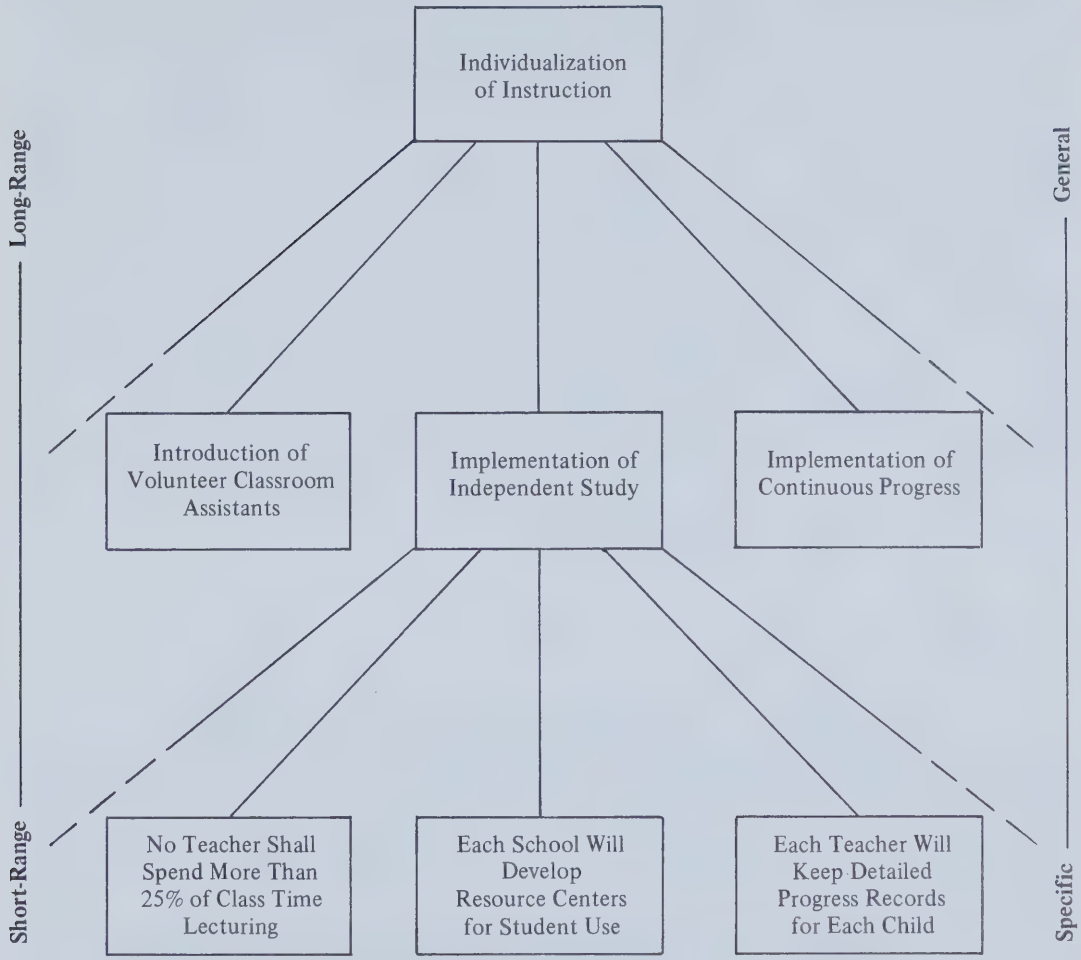
The purpose of defining enabling objectives centers around the role of teacher as facilitator of student learning. Enabling objectives are used in Procedures Evaluation, to ensure that the program is being implemented in a fashion that is consistent with the plans. Enabling objectives relate to all aspects of the process of instruction. Examples are facilities, grouping of students, materials, as well as teacher-learner transactions.

An example of an enabling objective is shown in Figure 3.1. This figure illustrates the various levels of an enabling objective.

Terminal objectives are expressions of the desired results of a program. Expected student behavior can be categorized as cognitive, affective, and/or psychomotor. Behavior expressed as a cognitive objective relates to specific knowledges and reasoning skills which the student is expected to master. An example of a cognitive objective is: *80% of the students shall state at least four of the causes of World War I.* This objective relates to the capability of the student to recall information. Another example would be: *Given knowledge of the causes of World War I, and World War II, the student will outline five current world problems which, if unchecked, could lead to a World War III.* This objective relates to the student's reasoning ability and the ability to transfer knowledge to a new situation. Objectives in the cognitive domain should go beyond straight recall and recognition if the student is to be prepared adequately for the changing society in which he will live.

The second area, the affective domain, relates to the child's ability to value. Interest, morale, self-image, enjoyment, and feelings are all part of the affective domain. Definition of affective objectives is a much more difficult task than definition of cognitive or psychomotor objectives. However, with curriculum emphasis being placed more and more in this area, the need for definition of such objectives is becoming more important. If curriculum which emphasizes values is to be evaluated, then affective objectives must be specified. The problem of evaluation in this area is related to the invisibility of values. The assumption must be made that if a child behaves in a certain fashion that he therefore holds certain values. Only the behavior can be observed and evaluated. The short-range terminal objectives must be specified in terms of observable characteristics; the corresponding long-range objectives outline the values which the teacher wishes the student to consider. An example of a short-range affective objective would be: *The student will, voluntarily, read six full-length novels during one semester.* The corresponding long-range objective would be: *The student will develop a love of good books and appreciation for English literature.*

Figure 3.1 Levels of an Enabling Objective


Individualization of Instruction
Long-Range
General
Introduction of Volunteer Classroom Assistants
Implementation of Independent Study
Implementation of Continuous Progress
Short-Range
Specific
No Teacher Shall Spend More Than 25% of Class Time Lecturing
Each School Will Develop Resource Centers for Student Use
Each Teacher Will Keep Detailed Progress Records for Each Child

The third area, the psychomotor domain, refers to the physical dexterity and coordination of the child. An example of a psychomotor objective would be: *At the end of Typing 10, 80% of the students shall be able to type, on the average, 50 words per minute or better, as shown by three one-minute speed typing tests.*

It must be recognized that every curricular area has some components of each of the three areas. The emphasis among these, however, may change over time. For example, in language arts, in the earliest years, more emphasis is placed on the physical coordination involved in being able to write, cut, and manipulate objects than is placed on these objectives in junior or senior high school. Even in areas that seem to be mainly one of these three major areas, aspects of the remaining two invariably enter. For example, in physical education, the majority of teaching time is spent in the psychomotor domain. However, throughout is the emphasis on teaching the child to enjoy various activities in order that he will use these skills for leisure activities in later years. This, of course, implies physical education objectives in the affective domain.

It should be noted that priority of objectives does not dicate allocation of teaching time. Although the affective domain may have highest priority, there is not necessarily any contradiction in a teacher spending most time in the cognitive area. Often, in order to influence interests, one of the appropriate roles for a teacher is to ensure that the child is exposed to a wide variety of experiences which he enjoys and about which he *may* develop an interest. Some of these experiences may be in the cognitive area. For example, a child cannot develop a love of books unless he is able to read and understand a sentence. These cognitive skills enable the affective objectives to be achieved. The reverse is often assumed to be true as well. That is, students tend to work harder in subjects that they enjoy. Figure 3.2 illustrates the levels of terminal objectives.

2. Specific vs. General Objectives

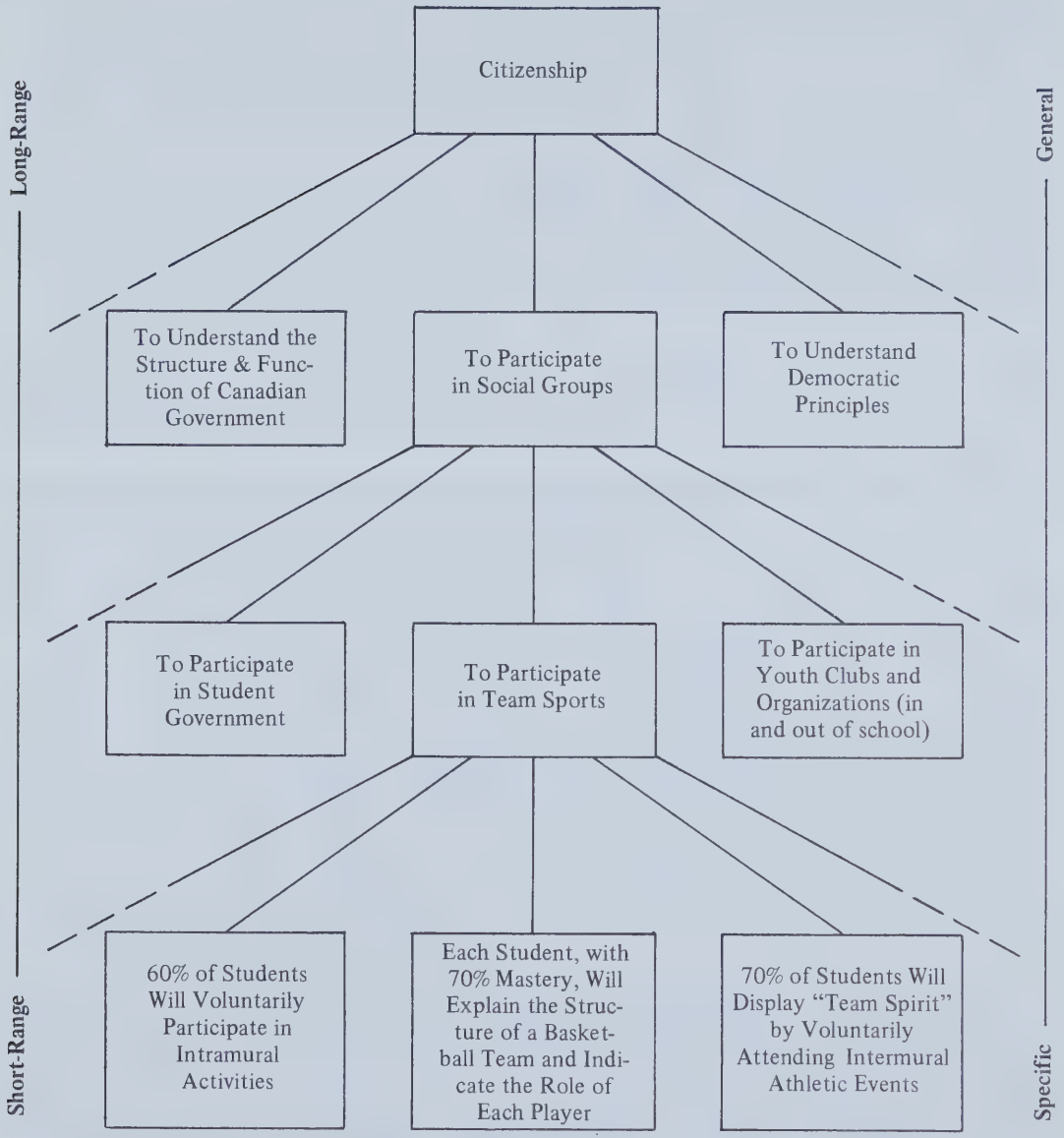
There are many levels of specificity of terminal and enabling objectives. There are societal goals, examples of which are the Aims and Objectives of elementary and secondary education. These are provincial goals which give a general direction to education. There are school-system goals, which are more specific and relate more to the needs of specific communities within the system. There are school goals, which are still more highly defined. Finally, there are classroom objectives which relate to a particular group of students and are the most highly defined.

It is possible that all school systems would have the same general terminal objectives. However, it is probable that the priorities between and among these objectives would differ. For example, an urban, inner city, high school may place the highest priority on objectives relating to employment and job opportunities for graduates. This would be reflected in an emphasis within the school on vocational, technical, and business education. On the other hand, a rural community with a large proportion of students for whom English is a second language may place highest priority on communication skills. This would be reflected in an extended language arts program.

General objectives are highly useful in planning an instructional program. However, to evaluate that program, more specific objectives are required.

The most specific form of objectives is often referred to as behavioral or performance objectives. Either terminal or enabling objectives can be put in this form.

Figure 3.2 Levels of a Terminal Objective


Citizenship
To Understand the Structure & Function of Canadian Government
To Participate in Social Groups
To Understand Democratic Principles
To Participate in Student Government
To Participate in Team Sports
To Participate in Youth Clubs and Organizations (in and out of school)
60% of Students Will Voluntarily Participate in Intramural Activities
Each Student, with 70% Mastery, Will Explain the Structure of a Basketball Team and Indicate the Role of Each Player
70% of Students Will Display "Team Spirit" by Voluntarily Attending Intermural Athletic Events
Long-Range
Short-Range
General
Specific

In order to evaluate a student's progress, a group of students, a curricular or extracurricular program, or an entire school, information is necessary. The basic source of such information is observed behavior. Behavioral objectives outline, in descriptive terms, the behaviors which one would expect if the underlying general objective has been achieved. If terminal objectives are stated in such a way that the instructional target is an observable pupil behavior, or an observable product of pupil behavior, then the quality of instruction can be determined by using these explicit criteria. Precision in statements of objectives is rare. Whereas general terminal objectives may be helpful to teachers in initially deciding on what they wish to teach, their vagueness decreases the potential assistance in guiding the teachers' selection of learning activities, or in deciding how to evaluate the worth of these learning activities. Only behaviorally stated terminal objectives allow for rigorous selection of learning activities and evaluation procedures.

However, stating a terminal objective in behavioral terms does not guarantee that the objective is meaningful or educationally valuable. Unfortunately, the types of pupil behavior that are most easily defined are often the most unimportant. Truly significant objectives are usually much more difficult to state in a specific fashion. For instance, it is much simpler to have a student recite multiplication tables than to demonstrate behaviors that reflect problem-solving ability. This, in fact, can be an advantage of defining specific objectives. When objectives are defined specifically, it becomes easier to decide toward which objectives one should teach. Use of criteria to judge the merit of terminal objectives requires that the objective be accurately interpreted. This can be accomplished by defining an objective in performance terms.

A behavioral terminal objective is stated in terms of the behavior that the student shows, which in turn reflects that he has achieved the objective. There is, however, another dimension to objective writing, a dimension which further aids the teacher in planning and evaluating instruction. It involves establishment of performance standards, that is, specifying, in advance of instruction, minimal levels of pupil achievement. As was detailed in Chapter II on Needs Assessment, a description of the present state or condition--a description of what is--allows the objective writer to develop realistic terminal objectives. For example, the objective, *The student will solve problems relating length, breadth, area and perimeter of a rectangle*, does not have a performance standard. Baseline, descriptive information of present achievement assists in developing a reasonable standard. If a student gets four out of ten problems correct, has he achieved the objective? The objective should be reworded to include such a standard, thus, *The student will solve six out of ten problems relating length, breadth, area and perimeter of a rectangle*. Now, suppose that ten children out of the class of 25 achieve this objective. Is this acceptable to the teacher? Does this warrant re-teaching of the material? It can be seen that if the teacher also wishes to judge his own performance in terms of the achievement of the total class, then a class minimal level is also required. The objective would then begin *70% of students . . .* or *Each student . . .*

Another example, this time in the affective domain, would read, *70% of students will attend, voluntarily, a free concert given by the local symphony orchestra.* Now, the general objective which is reflected in this specific objective relates to the establishment of an appreciation of good music. The behavior is voluntary attendance at a concert. The student minimal level is that he attend voluntarily, an either-or choice. In effect the student can achieve at a level of 100% or 0%, and only 100% is acceptable. The class minimal level is that 70% of the students do attend the concert.

It is important to remember that minimal levels should be established prior to the period of instruction being evaluated. (Descriptive baseline information from needs assessment

facilitates choice of minimal levels). In that way minimal levels can serve as standards by which the teacher can judge both his students and his own performance. If a teacher waits until after the instructional and testing period is over, and inspects the students responses without any previously established standards, it is probable that he will be more inclined to accept the results as the best that could be expected.

Similarly, in procedures evaluation specific enabling objectives facilitate evaluation.

Just as with terminal objectives, in order to improve the evaluation, enabling objectives should be based on the observable behavior of the instructional staff and those in supportive and administrative roles. An objective such as *The teacher will foster the maturity of the child* is not extremely helpful in deciding whether the teaching staff is succeeding in this worthy endeavor. However, an enabling objective which reads, *At least 60% of the time, the teacher will use questioning strategies which require more than just a memory response on the part of the student* allows behavior to be observed and therefore facilitates evaluation. Use of terminal and enabling objectives defined specifically greatly facilitates teaching, learning and the evaluation of both.

Figure 3.3 illustrates the type of objective used in various stages of evaluation.

Figure 3.3 Objectives Classification in Two Dimensions

	General	Specific
Terminal	– used in early Needs Assessment	– defined in late Needs Assessment; used in Outcome Evaluation
Enabling	– used in early Program Planning	– defined in late Program Planning; used in Procedures Evaluation

2. Short-range vs. Long-range Objectives

These levels of specificity can also indicate the immediacy with which the terminal or enabling objective will be met. It is usually the case that short-term objectives are met within one year or less. However, referring to Figure 3.2, the objective of producing a good citizen would probably take many years to achieve, and the result may not be apparent until years after the student has left the educational system. In other words, specific objectives are generally short-term in nature, whereas general objectives can be either short or long-term.

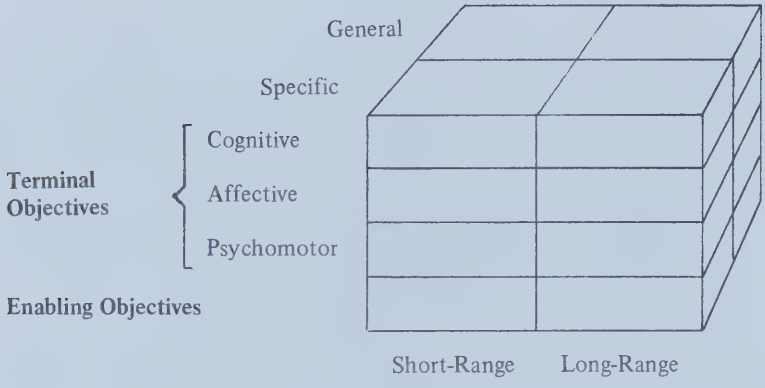
The majority of evaluation (which is always objective-oriented) is focused on the short-term objectives. In industry, the minute a car rolls off the assembly line, it can be

evaluated and can indicate, therefore, the worth of the production process. However, in education, the "product" may not show the effects of the educational system until many years later. Evaluation studies which follow a group of students throughout their careers are extremely expensive, and although there is a need for this type of information, educators are faced with pressing decisions about the worth of various programs in the here and now. These decisions cannot wait the twenty years needed to discover the achievement (or lack of achievement) of the long-range objectives of a program. Much of educational evaluation focuses, therefore, on the short-range, or at most medium-range objectives.

This discussion emphasizes the need for a rational hierarchy of objectives. Achievement of a short-term objective should move the students one step closer to achievement of the corresponding long-range objective. Evaluation based on short-range objectives is often generalized to the corresponding long-range objectives. But, this generalization cannot be made without a well-defined hierarchy of goals and objectives.

The three types of objective classification are illustrated in Figure 3.4. Use of this classification system aids the writing, and the evaluating of terminal and enabling objectives.

Figure 3.4 Objectives Classification in Three Dimensions



C. HOW TO WRITE TERMINAL OBJECTIVES

There are several characteristics of a "good" terminal objective. One sample list of such characteristics is given below:

1. The objective should be stated using the active form of the verb.
2. The objective should specify the behavior expected from each student.
3. The objective should specify the numbers or per cent of students expected to achieve this objective.

4. The objective should indicate how achievement of this objective will be evaluated.
5. The objective should indicate the length of time in which it is expected that this objective will be achieved.
6. The objective should indicate the conditions under which the behavior is expected to occur.

The usual approach to the writing of terminal objectives becomes almost a fill-in-the-blanks exercise which concentrates heavily on the structure of the objective. That is, a person writing an objective starts with a format such as:

WITHIN .... (time period) .... GIVEN .... (conditions) ...., ...(class minimal level) .... OF THE STUDENTS WILL .... (behavior-active verb) ....,AT A .... (student minimal level) .... LEVEL, AS EVALUATED BY .... (method of evaluation) .... .

and then proceeds to fill in the blanks.

It is felt that the major emphasis in writing terminal objectives should be placed on the content of the objective rather than on the structure of the objective. That is, *what* is said is more important than *how* it is said. There is nothing wrong with good structure for an objective, and the above "fill-in-the-blanks" is helpful in evaluating an objective. However, if priorities must be assigned in terms of time and resources, the content concern must be the crucial factor. Therefore, the approach taken to writing objectives in this publication reflects this content concern.

The purpose for defining and delineating a terminal objective is to communicate clearly one of the teacher's intended educational outcomes. Within the constraints of time and resources, there should be as many objectives as there are intended educational outcomes. If a written objective clearly defines a desired outcome, then the ultimate criteria, given below, can be answered in the affirmative.

THE ULTIMATE CRITERIA:

CAN ANOTHER COMPETENT PERSON SELECT SUCCESSFUL LEARNERS IN TERMS OF THE OBJECTIVE SO THAT YOU, THE OBJECTIVE WRITER, AGREE WITH THE SELECTIONS?

The method of writing terminal objectives which is given below is a cyclical or iterative procedure. That is, three steps are completed at each stage of definition, and then, if inconsistencies are found, the steps are repeated. These steps are outlined below:

1. Draft the rough objective.
2. Describe an evaluation technique or criteria which would indicate how successful completion of the objective would be indicated, e.g., a sample test item, or indicator.
3. Describe the level of pupil behavior which is called for, e.g., memory, problem-solving, valuing. (Bloom's and Krathwahl's Taxonomies can be helpful here.)

It is the third step of the cycle which points out the inconsistencies. That is, it is likely the draft objective calls for a high-level behavior, whereas the evaluative criteria is testing a much lower level. This would lead to either "beefing-up" the evaluative criteria or lowering the level of the objective.

Two examples are given to illustrate this process.

EXAMPLE: **A Cognitive Terminal Objective**

First draft objective: The student will learn to read.

Evaluative criteria: Observation of oral reading.

Performance level: Pronunciation (translation).

Comments

It seems that in order to read, the student should be able to pronounce words. This should be stated in the objective.

Second draft objectives: The student will learn to pronounce words orally from written work.

Evaluative criteria: Observation of oral reading.

Performance level: Pronunciation

The idea of comprehension seems to have been lost in this iteration. The objective should be expanded to allow for this factor. Also, a method of evaluating comprehension skill should be stated.

Third draft objective: The student will pronounce orally a sample of written work and answer questions on the content of the paragraph.

Evaluative criteria: Observation of oral reading. Test for content mastery.

Performance level: Pronunciation (translation) and comprehension.

It is at this point that structure enters the picture. Up to now, the concern has been content only. Now - how many words does the student have to read? What type of written work will he be given to read from? How many pronunciation errors are allowed? How much content should he absorb? The objective should outline these aspects so that observers other than the objective writer could choose the *same* successful learners.

Fourth draft objective: Given a paragraph of 100 words from a basic primary reader, the student will read it orally, making no more than five errors in pronunciation, and will subsequently answer correctly four out of six questions on the content of the paragraph.

Evaluative criteria: Observation of oral reading, test for content mastery.

Performance level: Pronunciation (translation) and comprehension.

Now, test the objective, Can another competent person select successful learners in terms of the objective so that you agree with the selections? The answer is yes, and we proceed to the next example.

It is also to be expected that much branching of objectives will occur. The usual procedure is to start at a position of a general objective. This general objective would lead to several more highly defined objectives which in turn may branch. By keeping track of these branches, the objective-writer facilitates the development of a complete set of objectives, all of which relate to the general goal. Figures 3.1 and 3.2, given earlier, illustrate this branching process.

EXAMPLE: **An Affective Terminal Objective**	Comments
First draft objective: The student will place a high value on education. **Evaluative criteria**: Average daily attendance. Dropout rate. Self-concept scale. **Performance**: Valuing	This is a general statement which leads to several specific objectives. One of these objectives, relating to the problem of school dropouts, is followed through below.
Second draft objective: The school's dropout rate will fall. **Evaluative criteria**: Dropout rate **Performance**: Acquiescence with values of society.	This is not used here as a terminal objective. Also, the performance indicated does not require the student to "value" education, merely to stay in school. In addition, with present legal requirements, many students cannot "drop-out" even though they would wish to do so.
Third draft objective: Fewer students will want to drop out of school. **Evaluative criteria**: Dropout rate; attitude scale measuring attitude toward school. **Performance**: Valuing	How much is "fewer?" What behavior does the student exhibit when he "wants to stay in school?"
Fourth draft objective: 80% of the students will achieve a score of 6.5 or better on the "Attitude Towards School" scale. No student shall score less than 5.0. **Evaluative criteria**: "Attitude Towards School" scale. **Performance**: Valuing	In this case, the behavior which is specified is operationally defined by the instrument which is used. The scale used in the example could be teacher-developed or a standardized test. It must, of course, validly measure whether a student has "psychologically dropped-out."

This example illustrates that the specified behavior could be achieving a certain score on a test. "Operationally defined" means that each item on the instrument need not be specified as a separate behavioral or performance objective. By examination of, and knowledge about the instrument one may imply the behaviors that are expected. In most cases, an "operationally defined" objective should relate to a standardized test or attitude scale rather than a teacher-made one, since the instruments developed by two different teachers may attempt to measure two different things. Teachers may not choose the same learners as successful since the two "teacher-made tests" will probably differ. The ultimate criteria for an objective does not hold, and the "teacher-made test" must be more fully defined.

D. EVALUATING THE TERMINAL OBJECTIVE

Objectives must be evaluated in order to ensure that they truly communicate the educational intent. The fill-in-the blanks format could be used for this purpose. Another method is shown, in checklist form, in Figure 3.5. There are two sets of columns, one of which relates to the adequacy of the objective (descriptive), and the other set which relates to the desirability of meeting the criteria. For example, an objective may have no time limit stated. In the case of individualized instruction, however, a time constraint may be undesirable. In addition, a few criteria for evaluating sets of terminal objectives are given.

E. HOW TO WRITE ENABLING OBJECTIVES

The procedure for writing enabling objectives is basically the same as for writing terminal objectives. To reiterate, the three steps are:

1. Draft the rough objective.
2. Describe an evaluation technique or criteria which would indicate how successful completion of the objective would be indicated.
3. Describe the type of behavior that is required to achieve the objective.

The content of an enabling objective should override the concern for format. That is, *what* is said is more important than *how* it is said.

The purpose of defining enabling objectives is to avoid ambiguity and to facilitate communication between and among the various persons involved in the program. Often, an instructional concept, such as team teaching, means slightly different things to different educators. By defining exactly what is meant by a procedure, through use of enabling objectives, all involved personnel are cognizant of what is expected of them.

An example of the process of defining an enabling objective follows:

EXAMPLE: **An Enabling Objective**	Comments
First draft objective: The teacher will facilitate student self-direction.	Here is another example of branching. There are many things a teacher can do to facilitate self-direction. One of these is followed through below.
Evaluative criteria: Amount of time which the teacher allows the student to take part in self-directed activity.	
Performance: Facilitative	

Figure 3.5 Evaluative Criteria for Terminal Objectives

	Description		Desirability	
	YES	NO	YES	NO
For Each Terminal Objective:				
(1) Is the objective trivial?				
(2) Does it describe, in clear terms, the behavior expected from the student?				
(3) Does it state a student minimal level of performance?				
(4) Does it state a class minimal level?				
(5) Does it outline the conditions under which the behavior is expected to occur?				
(6) Does it state a reliable way of evaluating the behavior?				
(7) Is there a time constraint?				
(8) Can another competent person select successful learners in terms of the objective so that you, the objective writer, agree with the selections?				
For Sets of Terminal Objectives:				
(1) Does the set of objectives cover all three domains?				
(2) Are the objectives all at a low level? Or do they include higher level objectives?				
(3) Are all the objectives short-term?				
(4) Do the specific objectives relate to the general objectives of the program?				

Second draft objective: The teacher will allow more time for student self-directed activity.

Evaluative criteria: Amount of time which the teacher allows the student to take part in self-directed activity.

Performance: Facilitative.

The content concern seems to be satisfied. But - how much is "more" time? What is a student self-directed activity?

Third draft objective: The teacher will spend no more than 20% of classroom time in large-group lecture situations, with the remainder of the time allocated to independent study and small-group discussion or project work, during which the teacher acts as a resource person.

Evaluative criteria: Observation of classroom activity.

Performance: Lecturing and acting as resource person.

Is the ultimate criteria met? Yes - other persons could choose which teachers are fulfilling the objective.

F. EVALUATION OF ENABLING OBJECTIVES

The "ultimate criteria" can again be put to good use in critiquing an enabling objective. That is, can another competent person decide if the objective has been achieved so that you, the objective writer, agree?

A suggestive checklist is given in Figure 3.6 for use in evaluating enabling objectives.

G. SUMMARY

The importance of well-stated terminal and enabling objectives cannot be understated. They provide most useful guidelines. In fact, the guiding principle of any evaluation should always be: refer to the objectives. During different stages of evaluation, different objectives are consulted. Needs Assessment and Outcomes Evaluation depend on the terminal objectives. Program Planning and Procedures Evaluation depend on the enabling objectives. In order for these objectives to be useful during the various stages, the objectives must be well-defined. A poorly-stated objective may be next to useless, since the prime purposes of removal of ambiguity and facilitation of communication will not be achieved. Evaluation of the terminal and enabling objectives, individually and in sets, will help to ensure that the evaluation's prime guidelines are capable of fulfilling their promise.

Figure 3.6 Evaluative Criteria for Enabling Objectives

	Description		Desirability	
	YES	NO	YES	NO
For Each Enabling Objective:				
(1) Is the objective trivial?				
(2) Does it describe, in clear terms, the behavior expected from the person(s) responsible for the achievement of the objective?				
(3) Does it state a minimal level of behavior?				
(4) Does it outline the conditions under which the behavior is expected to occur?				
(5) Does it state a reliable way of evaluating the behavior?				
(6) Is there a time constraint?				
(7) Can another competent person decide if the objective has been achieved so that you, the objective writer, agree with the selections?				
For Sets of Enabling Objectives:				
(1) Does the set of objectives cover the areas of administration, support, and instruction?*				
(2) Are all the objectives short-term?				
(3) Do the enabling objectives relate to the terminal objectives of the program?				

*Instruction includes most teaching activities; support includes activities performed by those who assist the teacher (teacher aide, guidance counsellor, librarian, etc.); administrative activities include such things as scheduling, grouping, budgeting, evaluating, etc.

REFERENCES CHAPTER III

EPIC Diversified Systems Corporation, *Developing and Writing Process Objectives*, Tucson, Arizona: Educational Innovators Press, 1971.

Kapfer, Miriam B. (editor), *Behavioral Objectives in Curriculum Development*, Englewood Cliffs, New Jersey: Educational Technology Publications, 1971.

Ramlow, Harold F., "Specifying useful instructional objectives" in Kapfer, Miriam B. (editor), *Behavioral Objectives in Curriculum Development*, Englewood Cliffs, New Jersey: Educational Technology Publications, 1971.

Alberta Department of Education, *Responding to Change: A Handbook for Teachers of Secondary Social Studies*, Edmonton: Department of Education, 1972, Chapter 1–3.

Bloom, Benjamin S. (editor), *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: The Cognitive Domain*, New York: David McKay Co., 1956.

Krathwohl, David R. and others, *Taxonomy of Educational Objectives. Handbook II: Affective Domain*, New York: David McKay, 1964.

Armstrong, Harold R., *A Teacher's Guide to Teaching-Performance Evaluation*. Worthington, Ohio: School Management Institute, 1972. (Address: School Management Institute, 6800 High Street, Worthington, Ohio, 43085).

IV
Instrumentation and Information

A. RELATIONSHIP BETWEEN OBJECTIVES AND INFORMATION

Evaluation has been defined as systematically collecting and processing information for the purpose of decision-making. The dual emphases in this definition are information and decision-making. Before decisions can be made, informaton must be made available in an understandable, relevant format. This chapter will attempt to delineate the concerns and cautions involved in collecting and processing information.

Once objectives have been defined, information must be sought which indicates whether the objectives have been achieved. A well-defined objective, whether terminal (student oriented) or enabling (procedures oriented), should indicate the method by which the stated behavior will be evaluated.

Early in the Needs Assessment stage of evaluation, it is probable that only general terminal objectives are in existence. By choosing a test, attitude scale or observation schedule to describe behavior, these terminal objectives are further defined so that finally, at the ends of Needs Assessment, they are explicitly stated in performance terms. For example, if the general objective relates to the self-actualization of each student, an appropriate scale to choose might be the Self-Concept as a Learner Scale (SCAL). In this case, self-actualization is operationally defined through the choice of instrument. The performance levels expected may be stated in terms of minimum levels of SCAL scores. On the other hand, there is no sense administering the Canadian Test of Basic Skills (CTBS) if the objectives of the evaluation relate to the peer group relationships of the student.

Two points should be noted:

1. Objectives define, implicitly or explicitly, the information required to allow an evaluator to judge whether the objectives are being achieved. The minimum expected performance levels included in the statements of the objectives provide the necessary standards.
2. Only information which is relevant to the objectives should be collected. Recording observations for the sake of observing is useless.

Information should have a purpose if it is to be gathered. If it is strongly felt that a certain test, attitude scale, or survey procedure should be utilized, and the resulting observations do not relate to the stated objectives of the program, then it is probable that the statement of objectives is not complete. Information collection must be related to objectives, since purposeless collection of information can be costly in terms of both time and resources and does not contribute to the evaluation.

B. CLASSIFICATION OF INFORMATION

Information can be classified according to three dimensions: (1) stage of evaluation; (2) source of instrument; and (3) type of information.

1. Stage of Evaluation Dimension

The stages of evaluation are Needs Assessment, Program Planning, Procedures Evaluation, and Outcome Evaluation. Each stage of evaluation requires various kinds of information. During the needs assessment phase, two basic types of information are required:

(a) descriptive, baseline information relating to "what is," and,

(b) normative, judgmental information relating to "what ought to be."

In a student-centered evaluation the information relating to "what is" describes the cognitive behavior or academic achievement, the affective behavior or attitudes and values and/or the psychomotor behavior. It is useful at this stage to describe "what is" both for the group that will be involved in the program, and also for other groups who will not be involved. This use of control groups greatly facilitates later decisions as to the worth of the program. The "what ought to be" is given by the statements of goals and objectives, which are also a type of information. Discrepancies which show up when these two types of data are compared indicate a need for action, either to modify "what ought to be" to bring it into line with reality or to attempt to improve "what is" so that it is consistent with "what ought to be." It may be that attention cannot be given to all the discrepancies which have been uncovered. A decision must therefore be made during the Needs Assessment stage regarding the selection of one or more problem areas. That information which has been collected relevant to the specific problem becomes an input to the Program Planning Stage.

During the Program Planning Stage the information that is required relates to alternative processes which will facilitate the achievement of the terminal objectives. Of course, one of the alternatives is the present method of instruction. Therefore, a description of the present procedures is a necessary type of information in this stage. Also, in order to describe other alternative processes, information regarding possible alternatives must be gathered. Some of the forms which this information may take are informed professional judgement, external sources such as the professional literature or reports from other, similar projects. In order to choose the best alternative, knowledge of constraints must be obtained. Such things as budget considerations, amount of relief time for teachers, and/or the availability of material and human resources that could be mobilized may be part of the necessary information. Once the choice of the "best" alternative has been made, a complete description of that alternative, in terms of enabling objectives, and including a budget, becomes an informational input into the Procedures Evaluation Stage.

During Procedures Evaluation, the enabling objectives, defined previously, become a standard for evaluation of the procedures which have been implemented according to the plan. In all probability the instruments (observation techniques) which would be used in Procedures Evaluation are the same as the instruments used earlier to describe and evaluate the old program prior to implementation of the new approach. In fact, use of the same instruments for both Program Planning and Procedures Evaluation aids the evaluator in making judgements about the comparative worth of the old and new programs. Costs of the two programs also give useful information about the implemented procedures.

Similarly, the same instruments or observation techniques could be used in the Outcomes Evaluation stage as were used to identify the cognitive, affective and psychomotor behavior of the students in the Needs Assessment stage. Again, this facilitates easier comparisons for indication of change. In order to ensure that the change is, in fact, due to the particular program being evaluated and not to maturation of the students, of non-program factors, one may use the same instruments in describing a similar group of students who are not involved in the program (control groups). The control group information assists

in controlling for many influences which may effect change in pupils. Control group information provides a standard on which to base judgements--the project group is expected to perform "better than" the comparison group.

In summary, the major considerations involved in the "Stage of Evaluation" dimension are:

(a) information gathered in each stage becomes an input for further stages;
(b) the number of *new* information gathering techniques (instruments) selected or oped for use in each stage decreases as the cycle is completed;
(c) similar types of information (resulting from use of the same instruments) are gathered in Needs Assessment and Outcomes Evaluation (product data) and in Program Planning and Procedures Evaluation (process data).

2. Type of Information Dimension

The types of information which can be used in an evaluation can be classified as:

(a) performance measures;
(b) judgmental responses;
(c) behavior observation;
(d) indicators, reports, documents;
(e) external information;
(f) other (budget data, constraints, objectives).

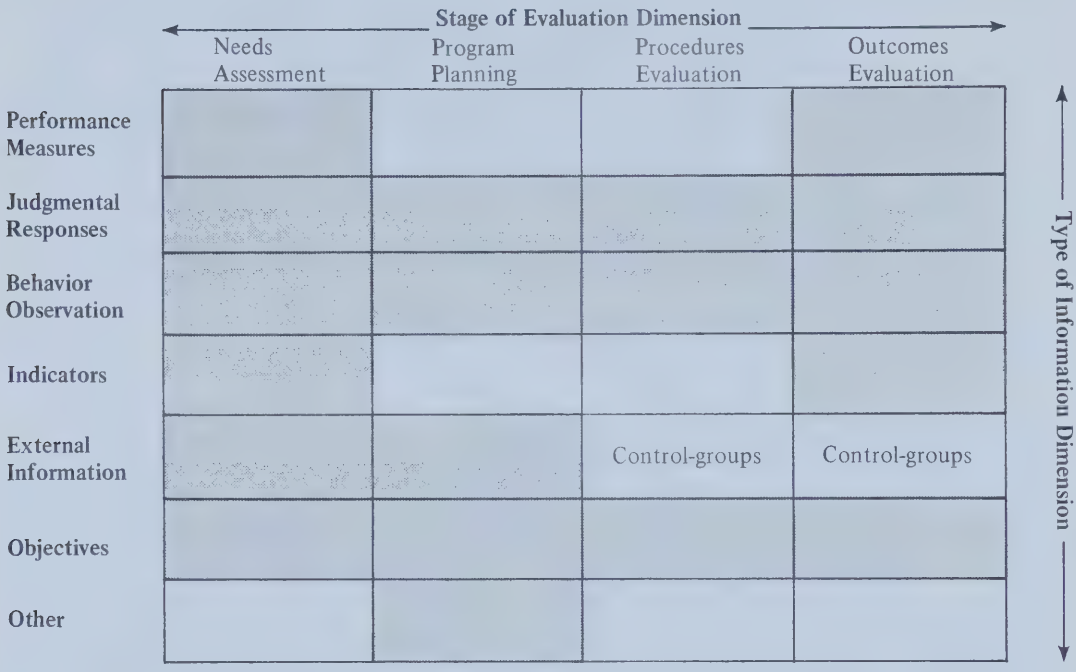
Each of these can be used at various stages of an evaluation. Those areas in which they seem to be most useful are indicated by the shaded areas in Figure 4.1. This does not mean that, for example, indicators should not be used during Procedures Evaluation, simply that it is likely that those areas which are shaded are commonly used in the illustrated stages. Each evaluation tends to be unique and the purpose of the evaluation as reflected in the objectives should be the only relevant criteria in choosing a particular type of instrument.

Performance measures are most often used for gathering information relative to terminal, cognitive objectives. Achievement and aptitude tests, work samples, and physical products are extremely useful in gauging achievement of such objectives. One advantage to using work products is that parameters and scoring rules can change after the students have done their work and without their involvement. For example, essays from two years of English may be compared even though the criteria were not made explicit until after the second year. Records of unit tests and end of year examinations also provide performance measures. In fact, the majority of the information developed in this category would probably be in the form of a written achievement test, whether teacher-made or published.

Classified as judgmental responses are such things as questionnaires, opinionnaires, surveys, interviews, and attitude scales. That is, relevant groups (students, teachers, parents, etc.) are asked to make a judgement relative to the area being evaluated. This is very useful information, although subjectivity can be a critical problem. Reliability is enhanced when the focus is the group and not the individual. However, given the state of the art, the informed judgement of competent individuals is probably the best single source of information for educational evaluation. Professional educators exercise a critical form of such judgement, that of deciding what information is relevant for an evaluation. One problem in collecting judgmental information is that of anonymity. That is, unless the respondents are assured of the anonymous nature of the information, responses may be biased.

Although the various judgements of relevant target groups is useful in every area of educational evaluation, most dependence is placed on this information in the evaluation of terminal, affective objectives and in the evaluation of enabling objectives.

Figure 4.1 Two Dimensions of the Classification of Information



Neither performance measures nor judgmental responses give the most accurate information in the evaluation of the instructional process. In this case, behavior observation is extremely helpful. However, this technique can be extremely costly in both effort and time. Observers must be well-trained so that the observations are as objective as possible. In addition, the fact that an observer is present may affect the behavior that is being observed. Even given these constraints, behavior observation can be the most revealing information obtained during an evaluation. Behavior observation provides a means to monitor, and then to describe, procedures. However, a qualitative evaluation of instructional processes is difficult. Standards are required. The only source of these standards is often the judgement exercised by the skilled, competent observer.

Indicators refer to a whole range of unobtrusive measures. An indicator can be defined as an indirect way of measuring behavior. For example, an indicator of increased respect for school and education may be a decreased amount of vandalism. Or, an improved valuation of education may be reflected in an improved average daily attendance. Vandalism or average daily attendance indicate that the desired attitude change is taking place. In the cognitive area an achievement indicator may be the number of scholarships received, or the number of students going on to higher education. It can be seen that indicators are usually very general, and could relate to a number of objectives.

External data is derived external to the system under consideration. This type of information normally plays a normative role in an evaluation. Such things as relevant studies, similar projects in other systems, groups of students from outside the system used for comparative purposes, all form part of this information category. With the exception of data obtained from external control groups, this type of information is most helpful during the program planning stage.

Information on objectives usually consists of statements of goals and objectives. These objectives may be locally produced or "borrowed" from published lists or from other jurisdictions. Objectives of some form are necessary throughout an evaluation. Enabling objectives are most useful during the Program Planning and Procedures Evaluation stages, whereas terminal objectives are most useful during Needs Assessment and Outcome Evaluation.

Other types of information that are useful in an evaluation include budget considerations, actual program costs, space limitations, in-service requirements, and available materials. The majority of this information would be used in the Program Planning phase of the evaluation.

The first two dimensions of the classification of information are illustrated in Figure 4.1 above.

3. Source of Instrument Dimension

Instruments, in general, can be purchased or locally produced. When planning an evaluation, consideration should be given to published or standardized tests. Caution should be exercised to determine if these tests actually are relevant to the objectives of the evaluation. Usually these instruments have been carefully refined and have known validity and reliability for a given population. An average evaluation could not afford the time and resources to construct an instrument of comparable quality. In selecting a published test, there are several considerations:

(a) The evaluator must have a clear idea of what is to be measured and select only those tests which do, in fact, measure it.
(b) The norm group on which the test was standardized should be as similar as possible to the group for whom the test is intended.
(c) Practical problems such as face validity (will the person taking the test consider it relevant), cost, test administration, and scoring must be considered.
(d) Technical excellence considerations are important: consult the test manual and other sources for reliability and validity information. The reliability coefficients (kind and magnitude) are important. The types of validity documented–e.g., content, construct and concurrent–assist in judicious selection.

The first point and the matter of validity in (d) are of paramount importance. Unless a published instrument can be found which accurately represents some or all of the objectives being evaluated, it would be better to develop a scale which appears to more accurately measure the desired behavior. That is, while most published instruments are superior in quality to those an evaluator could construct himself, they are valuable only if used correctly and for the right purposes. It is much better to build a locally produced instrument, with due care taken, than to use a standardized test which is not suited to the purposes of the evaluation.

Before constructing a new instrument, check whether there is a standardized one which relates closely to the objectives of the evaluation. If no such test exists, little time has been lost. If such a test does exist, expenditure of time and resources has been reduced.

Published instruments are most useful in the following areas:

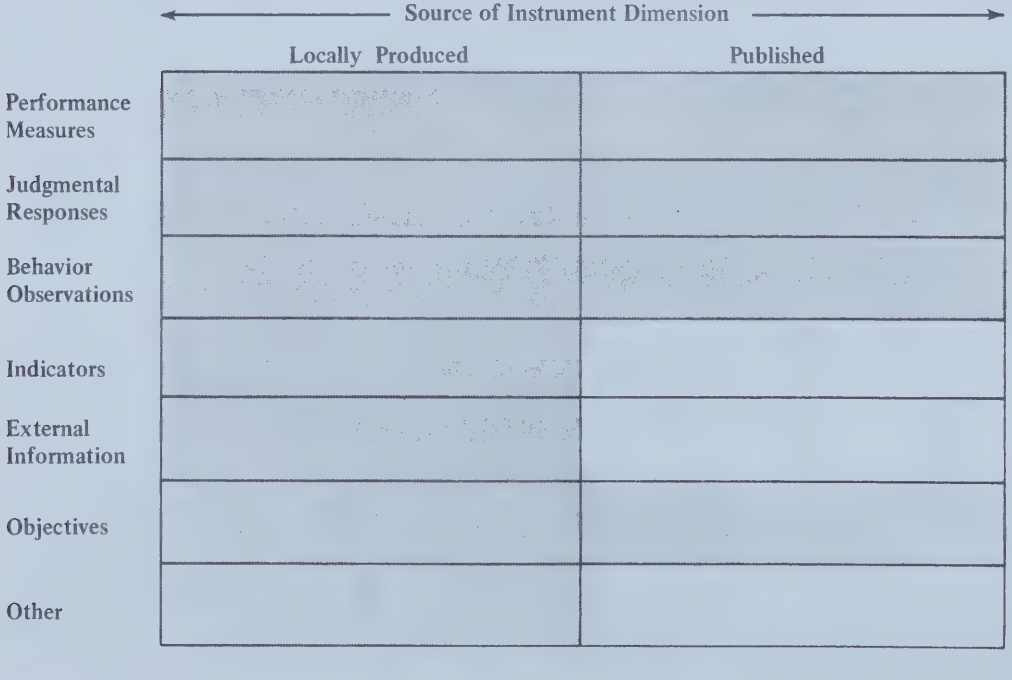
(a) Measurement of general global achievement in a cognitive area over a reasonably long period of time (one semester or grade). An example would be the California Reading Test.

(b) Measurement of a generalized affective component. An example would be the Self-Concept as a Learner Scale.

It is unlikely that a published instrument would be used to test achievement on a small unit of learning. For highly specific areas, it may be preferable, or necessary, to develop unique instruments. The next section indicates some general approaches to the development of instruments.

Figure 4.2 illustrates the usual sources of various types of instrument. Cognitive tests, or performance measures, are the most commonly used published instrument although several judgmental attitude scales are becoming more frequently used. Instruments used to catego-

Figure 4.2 Source of Instrument Dimension



rize behavior observations can be locally produced or published (for example, Flanders' Interaction Analysis Schedule). An anthology which describes literally dozens of observation schedules has been developed by Simon and Boyer (1967). Similarly, objectives can be locally derived or published lists of objectives can be used (for example, the Instructional Objectives Exchange, P.O. Box 24095, Los Angeles, California 90024, publishes sets of objectives with corresponding sample test items).

Figure 4.3 compiles the three dimensions of the classification of information.

C. HOW TO DEVELOP LOCALLY-PRODUCED INSTRUMENTS

All teachers have to make tests. However, these tests are usually in the cognitive domain. It is unusual to find teachers developing instruments which measure attitudes. Attitudes are inferred from observation of behavior. Although this professional judgement forms useful information, an independent examination of attitudes through a paper and pencil activity may provide an indirect means to measure them.

The generalized procedure for instrument development is the same for development of performance measures (cognitive domain) and judgmental responses (affective domain). There are basically five steps in this procedure.

Planning. The purpose of the test must be stated in terms of specific objectives (behavioral if possible). The test blueprint notion illustrated in Figure 4.5 is useful here. The blueprint format permits a check on the adequacy of the sampling of content and objectives.

Item Writing. The items must be carefully written and edited to ensure that they are measuring the stated objectives.

Try Out. The items must be assembled into a test and administered to a sample of subjects similar to those for whom the test is being designed.

Test Refinement. Information obtained from the try-out is used to refine the items and to adjust the administration and scoring procedures.

Final Test. After all corrections have been made, the test is ready to use.

Most teachers, although they do not follow this procedure explicitly, do in fact go through the above steps. Often, teachers have developed an item bank which relates to their objectives and which have been tested and refined over a period of years. These items are generally of two types: objective items (multiple-choice, true-false, matching) and essay items. The comparative advantages of these two items are outlined in Figure 4.4. Both of these types of items are useful in an evaluation of cognitive achievement. All items, it must be remembered, should relate directly to the objectives of the course. The correspondence between objectives and test items can be examined through the use of a format such as suggested in Figure 4.5. (Bloom's taxonomy is a useful way of classifying objectives). If it were discovered that a test covered only the lowest level objectives, it may indicate a need to add further questions. Similarly, a test may cover one unit of content very thoroughly, while omitting other important content areas.

In the affective domain, techniques of construction are not so well known, since few teachers administer, or develop, attitude scales or instruments requiring a judgmental response on the part of the person completing the instrument. In general, the procedure for developing an affective instrument is the same as for development of a cognitive test. However, item construction is slightly more complicated, since the background of knowledge is less developed. In order to develop an instrument that has face validity (i.e., the respondent feels that it is relevant), it is necessary to cover adequately the possible

Figure 4.3 Three Dimensions of the Classification of Information

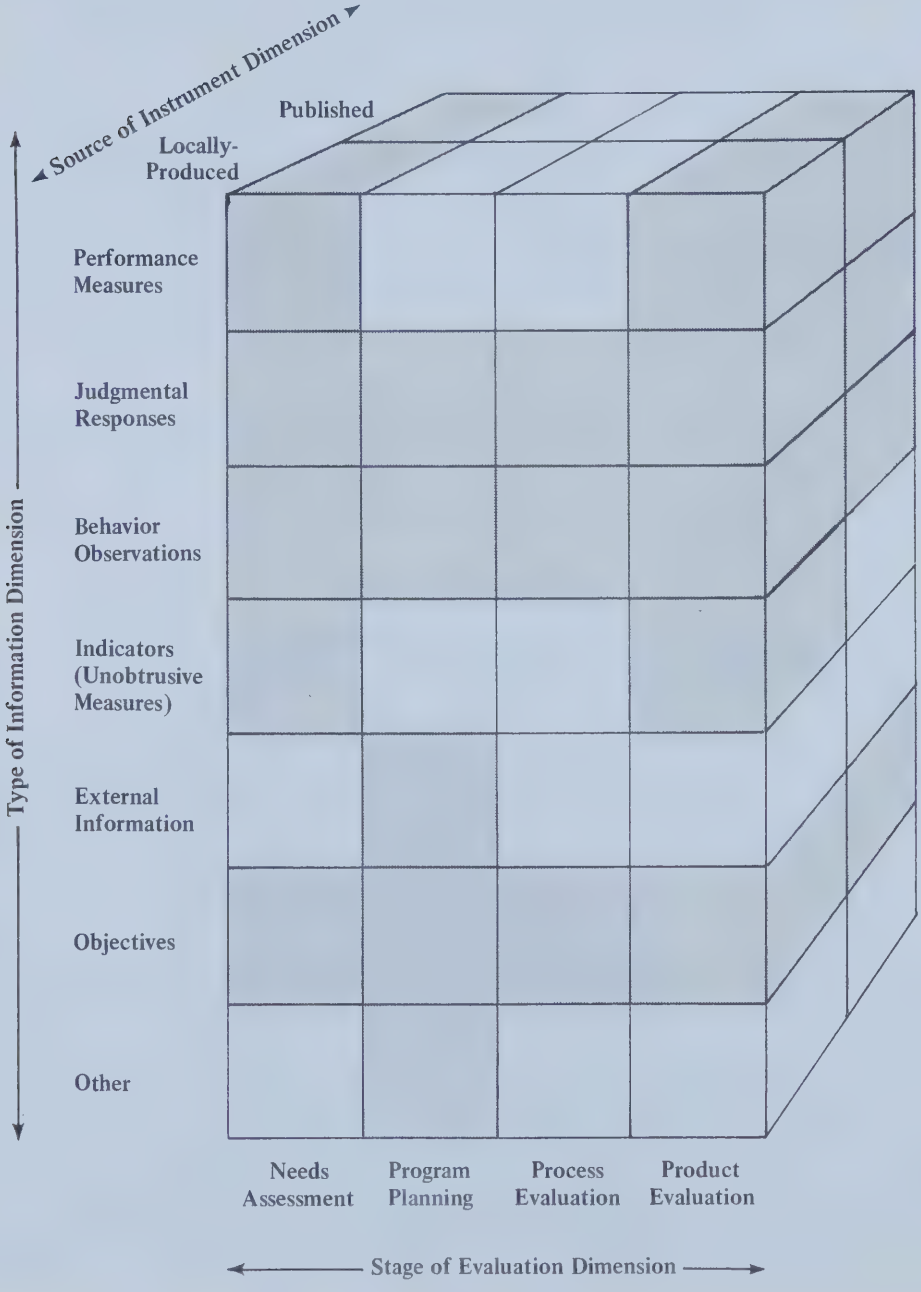


Figure 4.4 Essay Tests vs. Objective Tests*

	Advantages	Disadvantages
Essay Tests	- efficient for measuring ability to select and organize ideas, originality, writing ability. – only a few questions needed (relatively easy to prepare). – encourages concentration on larger units of subject matter.	– sampling of course content is limited. – enables bluffing and writing skill to influence score. – scoring is subjective, and slow.
Objective Tests	– efficient for measuring knowledge of facts. Some types can a lso measure understanding, thinking skills and other complex outcomes. – extensive sampling of course content. · scoring is objective, quick, easy, and consistant. – encourages concentration on detail, rather than on larger areas of content.	– large number of questions needed; preparation is difficult and time consuming. – subject to guessing.

* Adapted from Norman E. Gronlund, *Measurement and Evaluation in Teaching*, MacMillan Company, Collier-MacMillan Limited, London, 1969 pg. 108.

Figure 4.5 Test Blueprint – Classification of Test Items

TOPICS OR CONTENT AREA OF COURSE / LEVEL OF OBJECTIVE	UNIT I (Question No.)	UNIT II (Question No.)	UNIT III (Question No.)	...	TOTAL	%
KNOWLEDGE	1,18,35,37	8,25,28,38	6,7,27		11	27.5%
COMPREHENSION	34,40	10,17,26	11,24,31,32,39		10	25.0%
APPLICATION	9,19,20	4,16	2,30		7	17.5%
ANALYSIS	12,21	29	14,23		5	12.5%
SYNTHESIS	36	3,33	15		4	10.0%
EVALUATION	22	13	5		3	7.5%
TOTAL	13	13	14		40	100.0%

statements which suggest certain attitudes. One method of achieving this is to outline open-ended questions that relate to the objectives in question. Then, interview a small sample of the target group, carefully compiling the replies. Alternatively, or concurrently, replies to open-ended questions may be written by a sample of respondents. The vocabulary and phrasing to be used in the non open-ended instrument--the objective one–is checked against this experience. Rewrite these replies following one of the standard formats (given below) and test again to ensure that inconsistencies, vagueness, vocabulary, etc., are adequate to ensure unambiguity of each item. Then, the instrument can be piloted, revised again and administered to the appropriate group. This technique is a necessity in development of any instrument which purports to elicit judgmental responses from which conclusions can be drawn.

A number of formats are available for item construction. The most common forms are:
(a) rating scales (graphic or numeric);
(b) multiple-choice items;
(c) Likert scale;
(d) Semantic Differential.

When using a rating scale, the respondent selects the descriptor that best describes the object he is rating. The major difference between graphic and numeric rating scales is that the graphic scale allows a continuous choice, whereas in a numeric scale, the respondent has a given number of choices. The numeric scale consequently is easier to score. Examples of these two types of scales are shown below.

Most people think I am:

	1	2	3	4	5
NUMERIC RATING SCALE	a genius	intelligent	average	stupid	a moron
GRAPHIC RATING SCALE	a genius	intelligent	average	stupid	a moron

Multiple-choice items do not show strength of feeling as some other formats do. The example below indicates the type of judgment most amenable to this format.

My favorite subject is:

1) **English**
2) **Mathematics**
3) **Science**
4) **Social Studies**

Care must be taken to include all possible alternates in such a question. In addition, a simple ranking is not obtained. For example, if Mathematics were the response, it may be only marginally preferred to some other alternate, or it may be immeasurably superior. In addition, the marker has no idea of the ranking, in the respondent's eyes, of the other three alternatives.

Use of a Likert scale allows strength of feeling to be shown. The stem should be positive or negative, never a neutral statement. The five categories of response are usually SA-strongly agree; A-agree; U-unsure; D-disagree; SD-strongly disagree. For a desired positive response, the scoring is SA-5; A-4; U-3; D-2; SD-1. For a desired negative response, the scoring is SA-1; A-2; U-3; D-4; SD-5. Examples of both positive and negative statements are given below:

I would rather be at school than anywhere else	SA (5)	A (4)	U (3)	D (2)	SD (1)
Homework is a complete waste of time	SA (1)	A (2)	U (3)	D (4)	SD (5)

In order to avoid respondents establishing a response set, the items should alternate in a random fashion. This does complicate the scoring somewhat. This type of scale can be adapted for very young children in the following manner:

When I meet new people, I feel like this (circle one)

The teacher would read the item, and the children would circle the appropriate face.

The Semantic Differential is another technique which can be used. It is similar to the rating scales illustrated above, however, the format differs somewhat. Using adjectives which are polar opposites, the respondents (adults and older students) are asked to rate their attitude toward various factors on a seven-point scale, as shown below:

Attitude object – school

useful	___ 1	___ 2	___ 3	___ 4	___ 5	___ 6	___ 7	useless
good	___ 1	___ 2	___ 3	___ 4	___ 5	___ 6	___ 7	bad

For younger students, it is recommended that a five-point scale be used, with verbal cues, as shown below:

Attitude object – father

very kind	kind	average	cruel	very cruel
very wise	wise	average	foolish	very foolish

The major difference between rating scales and Semantic Differentials is that the semantic differential must use bi-polar (opposite) adjectives whereas a rating scale need not. The selection of bi-polar adjectives (or their synonyms) should be from lists provided by Osgood and others (1957) in the volume entitled *Measurement of Meaning*. Those that Osgood has classified as "evaluative" should be used to measure attitudes. Construction, administration and scoring of Semantic Differential scales are described in the same volume.

Just as in the cognitive area, where a test may consist of multiple-choice, short-answer, and essay questions, an instrument which requires a judgmental response may be made up of rating scales, Likert scales, and open-ended questions. The most important thing to remember is that, whatever format is chosen, the instrument must relate to the stated objectives of the program.

The examples given above have all been relevant to either Needs Assessment or Outcome Evaluation, that is, to terminal objectives. The same techniques can be used in describing the process of interaction and/or instruction, as reflected in the enabling objectives. A number of examples are given below:

MULTIPLE CHOICE – When I am teaching,
a) the students are very attentive
b) there is some disturbance
c) there is a great deal of disturbance

SHORT ANSWER Give the proportion of time spent in the following activities

1. lecturing _____ %
2. large-group discussion _____ %
3. small-group discussion _____ %
4. small-group projects _____ %
5. individual work _____ %

RATING SCALE My lessons are well-prepared:

all of the time	most of the time	some of the time	never

LIKERT SCALE The amount of materials available to my class is unsatisfactory

SA	A	U	D	SD

SEMANTIC DIFFERENTIAL

Attitude object – Principal

beneficial	1	2	3	4	5	6	7	harmful
wise	1	2	3	4	5	6	7	foolish

ESSAY-TYPE Describe, in detail, the organization of your classroom. Include physical facilities, pupil grouping, seating arrangements, etc.

The type of judgement required may dictate the choice of item format, since some of these techniques are more useful in obtaining certain types of information. The Semantic Differential technique, provided the scales are developed with care, has recognized validity as a generalized scale. If appropriate bi-polar adjectives are used the attitude toward any object may be measured. Thus, the same attitude scale can be used in different situations.

D. COLLECTION AND PROCESSING OF INFORMATION

When designing an instrument it is imperative to keep in mind the target group towards which it is directed. There are basically three target groups which may furnish useful information for an internal school evaluation: students, teachers, and parents. If a system-wide evaluation is planned, central office staff and the board of trustees must also

contribute to the evaluation. It is unusual to have an evaluation in which only one group furnishes information, the usual case being two groups, students and teachers. It is preferable that the students to be surveyed be of two types; those students that are involved in the program, and a control group of students who are not involved in the program. Students are the only group which can furnish information relative to the achievement of terminal objectives, although the judgement of professional educators and parents forms useful subsidiary data. If the student does not demonstrate, *in some way*, his achievement of certain outcomes, then the professional has nothing on which to base his judgement, and it becomes merely unsupported opinion. This is not say that students must complete objective tests in order that an evaluation can take place, rather that he is the only one who can indicate, through *some* behavior, that he has achieved (or not achieved) a particular objective. It is on his behavior that the teacher's professional judgement must be based.

Given time and resources, the more information, relative to the objectives, that can be obtained, the more reliable the conclusions of evaluation are likely to be. This implies that the more target groups that can be sampled the better. For example, in procedures evaluation, often the teacher is the only source of information. He is usually asked to make certain judgements regarding the teaching-learning process in his classroom. Suppose that in response to one of the examples in the previous section, the following information was obtained.

Give the proportion of time your class spends in the following activities:

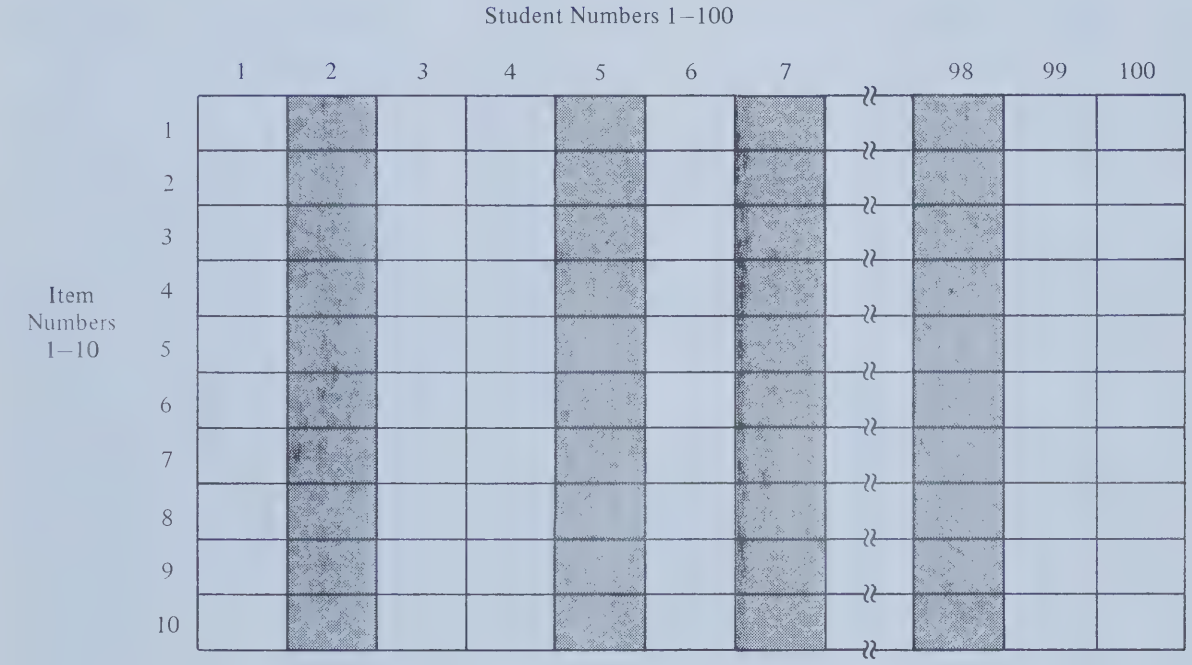
	Teacher Response	Student Response
1. lecturing	20%	70%
2. large-group discussion	10%	5%
3. small-group discussion	15%	5%
4. small-group projects	20%	10%
5. individual work	35%	10%

Normally, only the teacher furnishes such information. If only the teacher response were considered, the conclusion may be that the teacher had sufficiently individualized instruction and needs no further in-service. However, by also examining student response to the same question, different conclusions may be drawn.

It must be remembered here that it is what the student perceives is happening to him that influences his behavior, even if that perception is incorrect.

Few evaluations can afford to exhaust all information sources. The use of sampling techniques can ameliorate this situation. Briefly, sampling is a method of obtaining reliable information over respondents, over items, and/or over time. An example of sampling over respondents would mean that instead of gathering information from all students, only a proportion of students would have various instruments administered to them. If the sample is to be assumed to be representative of the total group (all pupils had an equal opportunity to be included) then the results may be generalized to the original population. If this group is a random sample, the information, in terms of averages, will be–except for sampling error--identical to that received from all students. Of course, sampling over students is useful only if the results of the testing program are to be used for evaluation of the program. If the purpose of administering various instruments is for individual pupil diagnosis then sampling is inappropriate. Figure 4.6 illustrates respondent sampling.

Figure 4.6 Respondent Sampling



Sampling over items avoids the problem of fatigue. That is, if the amount of information required from any one target group is large, this implies a lengthy instrument. Particularly when a group is requested to fill out the instrument on their own time and return it, a lengthy instrument causes the return rate to fall or perhaps too many respondents will not finish. A long instrument can usually be broken down to a number of shorter instruments, each of which contains incomplete information. However, by asking portions of the target group to complete different sub-tests, each being a part of the original instrument, the length of the task is decreased. This possibility is illustrated in Figure 4.7.

Whenever teachers develop an achievement test they are, in fact, selecting items from a large group of possible items. The sample selected from the teacher's item bank or repertoire is felt to represent the total area of concern for this particular evaluation of pupil progress. Figure 4.8 illustrates this form of item sampling.

Sampling over time is a slightly different procedure. Suppose that an evaluation was taking place concerning the use and function of a school library. It was decided to observe pupil behavior in the library. Since it would be possible to get quite different observations at different points in time, rather than choose one two-hour period for observation, more accurate data would be obtained by choosing twenty-four, five-minute periods scattered throughout the week, and, if possible, the semester. Student behavior would probably differ between Wednesday morning at 10:00 and Friday afternoon at 3:00; similarly, the first week of school would elicit different behavior from the week just prior to final exams. All these various types of behavior should be taken into account in the evaluation, and can be, through use of sampling over time.

One consideration that must be made in designing an evaluative instrument is the type of processing required. If an instrument contains essay-type or open-ended response questions then processing is extremely difficult and the compiled results are more likely to be subject to scorer bias. Objective items, using techniques such as rating scales, true-false, or multiplechoice, are amenable to computer processing, and allow rapid and accurate compilation. Although processing concerns should not be of prime importance in deciding on an instrument, processing must be considered. If it is not a consideration, one is likely to end up with masses of information which cannot be processed and compiled into a form which is useful to the decision-makers.

Statistics provide a useful tool in processing and reporting of information. However, advanced statistics, although needed to examine relationships and to indicate the probability that differences are due to chance, may not be necessary for a particular evaluation. In many cases, use of averages, medians, and ranges would be the most advanced calculations necessary. It must be noted that, by deriving a statistic, often it is impossible to translate this processed information back into its raw form. If the raw form of the information would be useful to the evaluation, care must be taken during processing that this information is not lost. All anticipated computations should be described during the planning of the evaluation so that the procedures for collection provide data which meet the assumptions necessary for completing and interpreting those computations.

Processing procedures must be flexible enough to take into account unexpected information, but unplanned-for information should be critically scrutinized to be sure that it is a valid indicator of needed information.

In summary, when collecting and processing information, remember that the fact that an item or observation was not originally intended to provide data is irrelevant. The more information used during an evaluation, and the greater its variety, the more likely that good educational decisions will result.

Figure 4.7 Respondent-Subtest Sampling

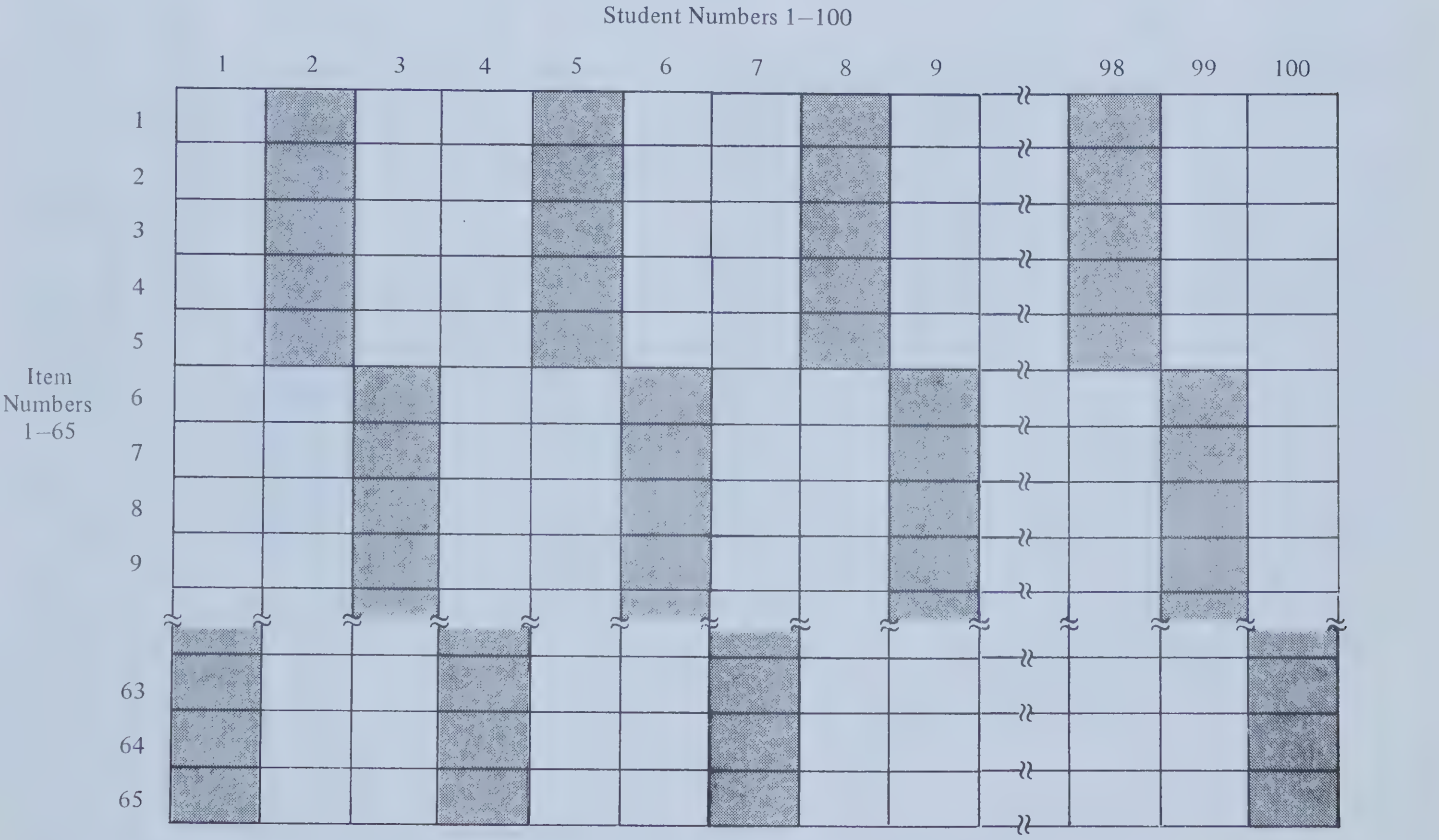
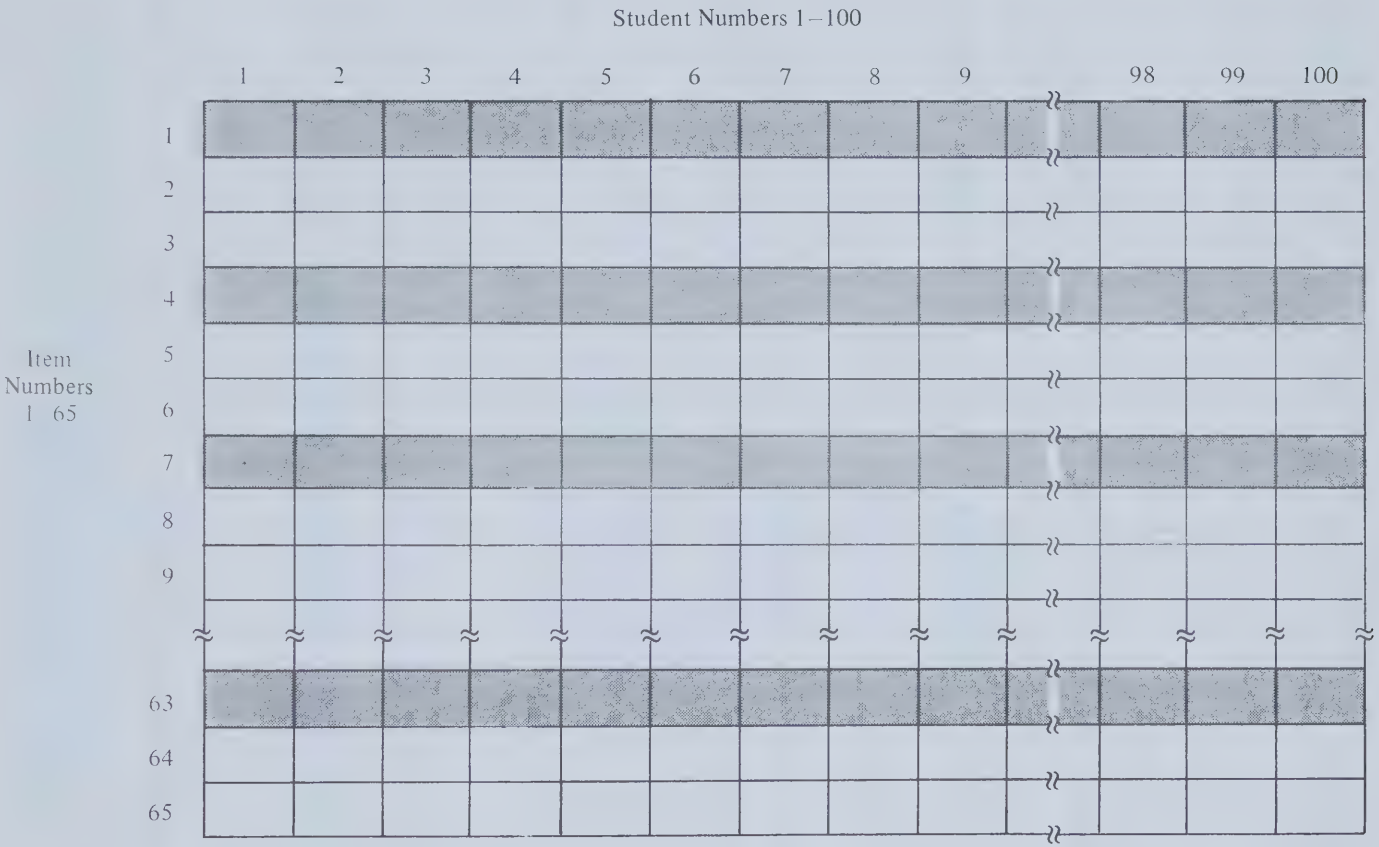


Figure 4.8 Item Sampling



E. EXAMPLE OF AN INFORMATION COLLECTION SCHEDULE

The following example outlines some of the information which may be needed in an evaluation. Due to space requirements, in most cases, only a portion of each type of information is shown. The checklists derived in Chapter II are used for format.

Needs Assessment

1. Statement of goals

(a) Each student should develop attitudes, skills and knowledge essential for effective living in a changing world.
(b) Each student should develop confidence in his own abilities, his sense of responsibility and his satisfaction and enjoyment of life.
(c) Etc.

2. Statement of General Objectives

(a) Each student should become proficient in thinking independently towards practical solutions to problems.
(b) Each student should acquire skills in information retrieval.
(c) Each student should be self-directed in the learning process, thereby accepting greater responsibility for his own learning.
(d) Each student should feel that he is a worthwhile person, and should develop a realistic appraisal of his strengths and weaknesses.
(e) Etc.

3. Description of Present Achievement in the Cognitive Area

(a) Canadian Test of Basic Skills (CTBS) results.
(b) Student files on past achievement on teacher-made tests.
(c) Etc.

4. Description of Present Attitudes

(a) Self-concept as a learner Scale.
(b) Study Habits Inventory.
(c) Teacher observation of students.
(d) Etc.

5. **Description of Discrepancies Between "What is" and "What ought to be"**

   (a) Results on CTBS Battery at or above Western Canada Regional norms, except in the area of math.
   (b) Problem-solving ability seems weak.
   (c) Low self-concept as a learner.
   (d) Little self direction evident.
   (e) Etc.

Decision: Program Planning will focus on the largest discrepancies, i.e., low achievement in math, low self-concept, low self-direction.

6. **Statement of terminal objectives**

   (a) Grade 8 students will increase their competence in math as measured by achieving a class average at or above Western Canada Regional Norms on the CTBS Modern Math Supplement.
   (b) Each student will raise his self-concept, as measured by the Self-Concept as a Learner Scale, by at least five points on the scale.
   (c) Each student will become more self-directed as measured by teacher rating of behavior.
   (d) Etc.

Program Planning Stage

1. **Description of Present Instructional Process**

   (a) Teacher questionnaire.
   (b) Student questionnaire.
   (c) Classroom observation.
   (d) Survey of library use.
   (e) Etc.

2. **Development of Alternative Processes**

   (a) Professional judgement that the present program is not adequate.
   (b) Visit to Brownville's Independent Study project.
   (c) Visit to Father Lacombe High School's Individualized Instruction Project.
   (d) Etc.

3. **Description of Factors Affecting Choice of Plan**

(a) Budget - School Board is willing to release an additional $5.00 per student for improvement of program.
(b) Teachers are willing to spend extra time in development of teacher-prepared materials.
(c) Teachers will require In-service to the extent of:
Individualized Instruction - 5 hours
Independent Study - 6 hours
Continuation of present programs - 1 hour.
(d) Etc.

4. **Description of Viable Alternatives**

(a) Individualized instruction.
(b) Independent study.
(c) Continuation of present program.
(d) Etc.

Decision: Independent Study will be implemented.

5. **Description of Proposed Plan**
(a) Teachers will allocate time on the following basis:

lecturing	20%
large-group discussion	10%
small-group discussion	15%
small-group project	10%
independent work	40%

(b) Separate resource center for math will be established.
(c) Teachers will develop Independent Study Units (I.S.U.) for student use.
(d) Etc.

6. **Statement of Enabling Objectives**

(a) Classroom 12 will be renovated to form a Mathematics Resource Center.
(b) One aide will be employed to man this resource center.
(c) Students will have free access to the resource center during Mathematics periods, study periods, and after school.

(d) Teachers will develop satisfactory I.S. Units as rated by an I.S.U. Evaluation checklist.
(e) Etc.

Procedures Evaluation

1. **Description of Implemented Instructional Process (Same Instruments as Program Planning)**

   (a) Teacher questionnaire.
   (b) Student questionnaire.
   (c) Classroom observation.
   (d) Survey of library and resource center use.
   (e) Etc.

2. **Description of Discrepancies Between Observed Procedures and Enabling Objectives**

   (a) Teachers are spending up to 50% of their time lecturing.
   (b) Some I.S. Units are not adequate.
   (c) Etc.

Decision: Program should be continued but further in-service for teachers is required. Revision to Independent Study Units required.

Outcome Evaluation

1. **Description of Achievement in the Cognitive Area (Same Instruments as in Needs Assessment)**

   (a) CTBS - Modern Mathematics Supplement.
   (b) Student records on teacher-made Mathematics tests.
   (c) Etc.

2. **Description of Attitudes**

   (a) Self-Concept as a learner.
   (b) Study habits inventory.
   (c) Teacher observation of students.
   (d) Etc.

3. Description of Discrepancies Between Outcomes and Intents

(a) Mathematics achievement not significantly different from prior to implementation.
(b) Self-concept improved, on the average, six points on the SCAL scale.
(c) Large, significant improvement in study habits.
(d) Etc.

Decision: Program should be continued in Grade 8 and expanded to include Grade 9 Mathematics.

F. SUMMARY

One factor that all evaluations have in common is the need for information. Inadequate or inaccurate information can lead to educationally unsound decisions. Therefore, to improve decision-making, information must be systematically collected and compiled into a useful format. However, not just any information will do. The instruments designed to collect information must be directly related to the terminal and enabling objectives. Unless the information is relevant to the objectives, time and resources will be wasted, since non-objective-oriented information is of no use to the decision-makers.

REFERENCES: CHAPTER IV

ITEM-SAMPLING AND RESPONDENT-SAMPLING

Cahen, Leonard S. and others, "The estimation of mean achievement scores for schools by the item-sampling technique," *Educational and Psychological Measurement*, 30, 1970, 41-60.

Cook, Desmond L. and Daniel L. Stufflebeam, "Estimating Test norms from variable size item and examinee samples," *Educational and Psychological Measurement*, 27, 1970, 601-610.

Husek, T.R. and Ken Sirotnik, *Item Sampling in Educational Research*, (ERIC ED013975), February, 1967.

Knapp, Thomas R., "An Application of balanced incomplete block designs to the estimation of test norms," *Educational and Psychological Measurement*, 28, 1968, 265-272.

Lord, Frederick, M., "Estimating norms by item-sampling," *Educational and Psychological Measurement*, 22(2), 1962.

Plumlee, Lynnette B., "Estimating means and standard deviations from partial data - an empirical check on Lord's item sampling technique," *Educational and Psychological Measurement*, 24(3), 1964, 623-630.

Shoemaker, David M., "Allocation of items and examinees in estimating a norm distribution by item-sampling," *Journal of Educational Measurement*, 7(4), Winter, 1970, 255-262.

Shoemaker, David M., "Item-examinee procedures and associated standard errors in estimating test parameters, *Journal of Educational Measurement*, 7(4), Winter, 1970, 255-262.

DEVELOPMENT AND USE OF ATTITUDE SCALES

Alberta Department of Education, *Responding to Change: A Handbook for Teachers of Secondary Social Studies*, Edmonton: Department of Education, 1972, Chapter 8.

Dawes, Robyn M. *Fundamentals of Attitude Measurement*, Wiley, 1971.

Festinger, Leon and Daniel Kutz, *Research Methods in the Behavioral Sciences*, New York: the Dryden Press, 1953, 340-353.

Fishbien, M. (editor), *Readings in Attitude Theory and Measurement*, New York: Wiley, 1967.

Fox, Robert and others, *Diagnosing Classroom Learning Environments*, Chicago: Science Research Associates, Inc., 1966.

Jahoda, Marie and others, *Selected Techniques*. Part 2 of *Research Methods in Social Relations*, New York: The Dryden Press, 1951, 423-462. See also Part I pp. 151-176 and 295-304.

Oppenheim, A.N. *Questionnaire Design and Attitude Measurement*, New York: Basic Books Publishers, 1966.

Osgood, Charles E. and others, *The Measurement of Meaning*, Urbana, Ill.: University of Illinois Press, 1957.

Parten, Mildred, *Surveys Polls and Samples*, New York: Cooper Square Publishers, 1966, 157-217.

Sellitz, Claire and others, *Research Methods in Social Relations (Revised)*, New York: Holt, Rinehart and Winston, 1951, 243, 255, and 552-573.

Shaw, Marvin E. and Jack M. Wright, *Scales for the Measurement of Attitudes*, Toronto: McGraw-Hill Book Company, 1967.

Summers G.F., *Attitude Measurement*, Chicago: Rand McNally, 1970.

Torgerson, Warren S. *Theory and Methods of Scaling*, New York: Wiley, 1958.

MISCELLANEOUS: INFORMATION GATHERING AND PROCESSING

Bloom, Benjamin S. (editor), *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: The Cognitive Domain*, New York: David McKay Co., 1956.

Bloom, Benjamin S. *Handbook on Formative and Summative Evaluation of Student Learning*, Toronto: McGraw-Hill, 1971.

Buros, Oscar Krisen (editor), *The Mental Measurements Yearbook*, Highland Park, N.J.: Gryphon Press.

Campbell, Donald T. and Julian C. Stanley, *Experimental and Quasi-Experimental Designs for Research*, Chicago: Rand McNally, 1966.

Center for the Study of Evaluation, *CSE Elementary School Test Evaluations*, Los Angeles: Center for the Study of Evaluation, UCLA Graduate School of Education, 1970.

Dizney, Henry, *Classroom Evaluation for Teachers* Dubuque, Iowa: Wm. C. Brown Company Publishers, 1971.

EPIC Diversified Systems Corporation, *Developing and Writing Process Objectives*, Tucson: Educational Innovators Press, 1971.

Gage, Gerald, "Distribution of Information," in Casper F. Paulson Jr. (editor) *A Strategy For Evaluation Design*, Monmouth, Oregon: Teaching Research Division of Oregon State System of Higher Education, April, 1970.

Gronlund, Orval G. and James W. Bommarito, *Tests and Measurements in Child Development*, San Francisco: Jossey-Bass Inc., Publishers, 1971.

Krathwohl, David R. and others, *Taxonomy of Educational Objectives, Handbook II: Affective Domain*, New York: David McKay, 1964.

Lyons, Tom, "Instrumentation," in Casper F. Paulson Jr. (editor) *A Strategy for Evaluation Design*, Monmouth Oregon, Teaching Research Division of Oregon State System of Higher Education, April, 1970.

Medley, Donald M. and Harold E. Metzel, "Measuring classroom behavior," in N.L. Gage (editor) *Handbook of Research on Teaching*, Chicago: Rand McNally, 1963, 247-328.

Metfessel, Newton and William Michael, "A paradigm involving multiple criterion measures for the evaluation of effectiveness of school programs," *Educational and Psychological Measurement*, 27(4), Part II, 1967, 931-943.

Paulson, Casper F. Jr., "Evaluation of instructional system," in Jack Crawford (editor) *National Research Evaluation and Development Training (CORD) Manual*, Monmouth Oregon: Teaching Research Division, Oregon State System of Higher Education, 1969.

Paulson Casper F. Jr., *A Strategy for Evaluation Designn*, Monmouth, Oregon: Teaching Research Division of Oregon State System of Higher Education, 1970.

Simon, Anita (editor) and others, *Mirrors for Behavior: An Anthology of Classroom Observation Instruments*, Philadelphia: Research for Better Schools Inc., Temple University Center for the Study of Teaching, 1967.

Walter, Larry J., "Information processing," in Casper F. Paulson Jr. (editor) *A Strategy for Evaluation Design*, Monmouth, Oregon: Teaching Research Division of Oregon State System of Higher Education, 1970.

Webb, Eugene J. and others, *Unobtrusive Measures: Non-Reactive Research in the Social Sciences*, Chicago: Rand McNally, 1966.

V
Budgeting for the Program and Its Evaluation

If the program under examination is to be compared to alternative programs in terms of cost a record of expenditures is necessary. The same information is required if the educational benefits (objectives) achieved are to be related in any way to costs. Accordingly, if program expenditures are to be monitored and audited, budget considerations should be included in program planning.

If the program during its various stages is to be evaluated, then early consideration of the costs of evaluation is indicated. A budget for evaluation may be supplemental to the program budget mentioned in the preceding paragraph. Early consideration of the time, resources, materials and personnel required for the evaluation often is in contrast to general practice. Too often a program is well under way before the staff time and funds needed for the evaluation receive attention. The school system's budget may have been fixed for the current fiscal year which requires that personnal and funds be "borrowed" from other activities.

Several budgets may be required for the program and its evaluation; one for each of the stages through which the evaluation of an instructional program normally proceeds. The least funding is required if the decision is taken to discontinue the program during the needs assessment stage. If the activity proceeds into the program planning stage additional resources are required. Personnel time and funding must be augmented still further–perhaps substantially--if the decision is to implement the plan selected. The evaluation aspect of this third phase–procedures evaluation--may be expensive. Outcome evaluation may be less demanding of time and resources but at this final stage most of the expenditures have already been committed.

Well in advance of each stage of program development and evaluation the reader may budget for the resources required. The list below is suggestive only and is not meant to be exhaustive.

Additional purposes served by the budget include meeting requirements of the local school board, the provincial government or any other funding agency.

Adaptations of checklists provided in previous chapters may be utilized for purposes of budgeting as well as planning. Additional columns may be utilized to provide space to record estimated costs and man-days.

For a particular on-going program some readers will feel, using the four-stages format, that they are commencing evaluation at the procedures or outcome stages. For on-going programs this initial activity is actually a needs assessment. If recommendations are formulated which suggest plans (or the need to plan new procedures) then program planning commences. If follow-up activities continue the sequential stages of planning, procedures evaluation and outcome evaluation may describe the activities undertaken. However, if no follow-up occurs evaluation has actually been terminated at the needs assessment stage.

Figure 5.1 Budgeting for the Resources Required

Budget item \ Stage	Needs Assessment $ *	Program Planning $	Procedures Evaluation $	Outcome Evaluation $
1. Released time for school system personnel				
2. Inservice				
3. External consultants				
4. Intervisitation				
5. Library and references				
6. Development and/or purchase of instrumentation and observation schedules				
7. Collecting and recording information				
8. Processing information				
9. Summarizing and interpreting information				
10. Preparing evaluation report				
11. Program materials				
12. Space and facilities				

*For some items estimates of time may replace dollars.

VI
Sources of Assistance

Groups or individuals who wish to evaluate a present program or a new one may require assistance beyond what is immediately available. Such assistance may take a number of forms. A discussion of these forms, and the evaluation stage at which they are most needed, is the substance of this chapter.

An obvious source of assistance is an educator within the school system or from outside who possesses expertise relevant to the target program. Such a consultant may assist in clarifying the terminal objectives developed during needs assessment and used during product evaluation. Consultation from this source will also assist in Program Planning where enabling objectives are used to specify the ways and means thought to be most efficacious. Finally, if one purpose of the evaluation is to ascertain the degree to which enabling objectives have been achieved (Procedures Evaluation) the expert in methodology has much to contribute. In Procedures Evaluation the observations and judgments of competent educators may be the only source of the standards requisite to a qualitative assessment (quantitative and descriptive information on support, administrative and instructional procedures may often be obtained by other means).

Indices which may be used for locating helpful printed information relevant to the program were listed in Chapter II in the discussion on program planning. Such a review of experience elsewhere is obviously beneficial. New information on objectives and their measurement in addition to other facets of the program may be uncovered in appropriate articles.

Chapter II also suggested visits to similar programs. Observation of the dynamics of some else's program may minimize errors in the local one. As with the review of the literature the experience of others may effect economy of action at home.

The processing of information may pose a problem. Early consultation and planning is advised. If data is to be machine-processed, consultation with the data processing centre must occur before collection. This advice holds for all treatment of data. If answers to specific questions are sought--perhaps the statistical relationships between certain program factors–early consultation is mandatory. Information on published tests, scales and observation schedules may be located in the accompanying manuals. Other sources are the *Mental Measurements Yearbooks* and *Tests in Print* referenced in Chapter II. Selection of the appropriate procedure to describe the current program, develop objectives and to measure achievement of objectives is thus facilitated.

In summary, Figure 6.1 outlines common sources of assistance.

Specific sources of Consultant assistance are listed as follows:

1. The specialist teacher(s) or consultant(s) within the school system.
2. The specialist teacher(s) from other school systems.
3. The appropriate consultant(s) from the Regional Offices of the Department of Education.
4. Professional development consultants of the Alberta Teachers' Association.
5. Appropriate personnel from post-secondary institutions.

Figure 6.1 Common Sources of Assistance

Stage of Evaluation	Needs Assessment		Program Planning		Procedures Evaluation	Outcome Evaluation
Evaluation Activity:	Description of Actual Outcomes	Dev. Terminal Objectives	Description of Actual Processes	Dev. Enabling Objectives	Achievement of Enabling Objectives	Achievement of Terminal Objectives
Sources						
Consultants	X	X	X	X	X	X
Literature Review		X		X		
Visits to Programs		X		X		
Data Processing Centre	X		X		X	X
Technical Manuals for Specific Instruments and Observation Schedules	X	X	X	X	X	X

EXTRA COPIES
OF CHECKLISTS

EVALUATION CHECKLIST

Checklist used for (check one):

- Planning the evaluation design ☐
- Recording progress of the evaluation ☐
- Reviewing the evaluation ☐

Evaluation Activity or Evaluation Product	Person(s) responsible							Dates	
		x	?	-	0			Initiated	Completed
1. Philosophical goal statement(s) concerning program purposes									
2. Statements of general objectives									
3. Descriptions of present state of condition.									
4. Relating descriptive statements to what "ought to be" -- Statements of specific objectives.									
5. Statements of specific objectives in performance terms									
- terminal objectives									
- related enabling objectives									
6. Quality of information									
- relevance and validity									
- reliability									
- breadth and variety									
- credibility									
- accessibility									
7. Performance standards against which information is compared									
- absolute standards									
- relative standards (information from norm or control groups)									

Satisfactory
Indeterminate
Unsatisfactory
Not Required

	Person(s) responsible	+ Satisfactory	? Indeterminate	- Unsatisfactory	0 Not Required	I Incomplete	Dates Initiated	Dates Completed
8. Evaluation resources								
- budget prepared								
- funding arranged								
- coordinator appointed								
- responsibilities assigned								
- time and/or facilities arranged								
- objectives development								
- and these activities related to information:								
selecting procedures								
collecting								
processing								
summarizing								
interpreting								
recommending								
reporting								
9. Questions and concerns of decision-makers (list the appropriate groups) Group Question or concern								
10. Decisions taken regarding the program under examination								
- to continue and adopt								
- to modify								
- to terminate								
- to expand								

NEEDS ASSESSMENT CHECKLIST

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
1. Statement of goals						
2. Statement of general objectives						
3. Description of present achievement in the cognitive area Instruments/Indicators/ Observation schedules used						
4. Description of present attitudes (affective areas) Instruments/Indicators/ Observation schedules used						
5. Description of present psychomotor abilities Instruments/Indicators/ Observation schedules used						

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
6. Description of discrepancies between "what is" (cognitive, affective and psychomotor levels #3, #4 and #5 above) and "what ought to be" (general objectives)						
DECISION: Which particular discrepancies will become the focus of the program planning stage? 7. Statements of terminal objectives						
8. Statements of terminal objectives in performance terms						

PROGRAM PLANNING CHECKLIST

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
1. Description of present instructional process						
Instruments/Indicators/Observation schedules used						
2. Development of alternative processes or programs						
Review of the literature						
Visits to similar projects						
Judgments of professionals						
Other						
3. Description of factors affecting the choice of a plan						
Budget considerations						
Release time for personnel						
Readiness of professional staff						
Inservice requirements						
Materials available						
Other						

Continued . . .

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
4. Description of viable alternatives						
DECISION: Choice of preferred alternative proposal (most viable alternative in terms of benefits and costs)						
5. Description of proposed plan and budget						
6. Statement of enabling objectives*						
Administrative						
Support - personnel						
- facilities						
- materials						
Instructional						
7. Relate enabling objectives to relevant terminal objectives.						

*Instruction includes most teaching activities; support includes activities performed by those who assist the teacher (teacher aide, guidance counselor, librarian, etc.); administrative activities include such things as scheduling, grouping, budgeting, evaluating, etc.

PROCEDURES EVALUATION CHECKLIST

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
1. State the enabling objectives						
2. Relate (1) to relevant terminal objectives						
3. Observe (monitor) the implemented instructional procedures Instruments/Indicators/ Observation schedules used _____ _____ _____						
4. Describe (record) observations of the procedures						
5. Description of discrepancies between observed procedures and enabling objectives						
6. Apply standards (make judgments) concerning discrepancies						
7. Summarize information from 1 - 6						
DECISION: Should the program be continued? Terminated? Modified? 8. If the decision is to modify-- suggest revisions to enabling objectives (formulate recommendations)						
9. Prepare the report for the procedures evaluation.						

OUTCOME EVALUATION CHECKLIST

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
For each group of terminal objectives:						
1. Review the statements of terminal objectives						
2. Observe (monitor) the relevant outcome behaviors:						
of achievement in the cognitive area						
Instruments/Indicators/Observation schedules used						
_____						
_____						
_____						
of achievement in the affective area						
Instruments/Indicators/Observation schedules used						
_____						
_____						
_____						
of achievement in the psychomotor area						
Instruments/Indicators/Observation schedules used						
_____						
_____						
_____						

Continued . . .

	PERSON(S) RESPONSIBLE	SATISFACTORY	UNSATISFACTORY	NOT APPLICABLE	DATE INITIATED	DATE COMPLETED
3. Describe (record) observations of outcome behaviors						
4. Describe discrepancies between observed outcomes and the terminal objectives						
5. Apply standards (make judgments) concerning discrepancies						
6. Summarize information from 1 - 5						
DECISION: To terminate? To continue? To modify? 7. If decision is to modify suggest revisions to terminal objectives (formulate recommendations), or if information is available on the related enabling objectives suggest revisions (formulate recommendations). If information is unavailable re: enabling objectives initiate procedures evaluation						
8. Prepare the report for the outcome evaluation						